

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: February 21, 2002, 16:26:14 : Search time 12.75 Seconds  
(without alignments)  
131.438 Million cell updates/sec

Title: US-08-753-851-1  
Perfect score: 123  
Sequence: 1 CERNGRYSISRTEADCKAFN 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4240

Minimum DB seq length: 0  
Maximum DB seq length: 22

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	26.8	17	2 A34835	ribosomal protein
2	28	22.8	19	2 B39845	pyrB leader peptid
3	27	22.0	21	2 I54351	gene HEXA protein
4	26.5	21.5	19	2 S63435	manganese peroxida
5	26	21.1	11	2 T06383	hypothetical prote
6	25	20.3	14	2 PH1598	Ig H chain V-D-J r
7	25	20.3	17	2 A54534	heat-stable entero
8	24.5	19.9	19	2 S65434	manganese peroxida
9	24.5	19.9	20	2 S68222	defensin AMP2 - Da
10	24	19.5	15	2 D48394	major fat-globule
11	24	19.5	20	2 PQ0544	capsid protein VP5
12	24	19.5	20	2 S67990	neurotoxin-associa
13	24	19.5	20	2 A56894	intracystalline c
14	24	19.5	22	1 MXKN1	mu-conotoxin GIIIA
15	24	19.5	22	1 MXKN2	mu-conotoxin GIIIB
16	24	19.5	22	1 MXKN3	mu-conotoxin GIIIC
17	23.5	19.1	22	2 A41833	40K iron-repressed
18	23	18.7	9	2 PH0942	T-cell receptor be
19	23	18.7	11	2 PH1632	Ig H chain V-D-J r
20	23	18.7	13	2 A61210	antibiotic GE2270
21	23	18.7	15	2 C44101	calmodulin, vasoac
22	23	18.7	19	2 PC1309	small granule S2 c
23	23	18.7	20	2 PT0248	Ig heavy chain CDR
24	23	18.7	20	2 PS0087	microbial serine p
25	23	18.7	20	2 I54283	arylsulfatase A -
26	22	17.9	12	2 PH1611	Ig H chain V-D-J r
27	22	17.9	15	2 PH0779	T-cell receptor al
28	22	17.9	17	2 C37520	glutathione transf
29	22	17.9	17	2 A36727	cytochrome c551 -

30 22 17.9 17 2 A27486  
31 22 17.9 18 2 PH1621  
32 22 17.9 20 2 S50175  
33 22 17.9 21 2 C31182  
34 22 17.9 22 2 I37144  
35 21 17.1 7 2 S16364  
36 21 17.1 7 2 S16365  
37 21 17.1 12 2 PH0746  
38 21 17.1 14 2 A60622  
39 21 17.1 14 2 PH1601  
40 21 17.1 16 2 F58501  
41 21 17.1 17 2 A34572  
42 21 17.1 17 2 S66364  
43 21 17.1 18 2 G02018  
44 21 17.1 18 2 S70612  
45 21 17.1 19 2 A49254

follitropin inhibi  
Ig H chain V-D-J r  
kallikrein (PK-120  
hypothetical 2.4K  
aspartylglycosamin  
opacity protein P.  
opacity protein P.  
T-cell receptor be  
somatostatin - spo  
Ig H chain V-D-J r  
25K kidney and gal  
osteonectin - rat  
sodium-translocati  
proteasome chain L  
alpha-macroglobuli  
tcr delta chain V-

#### ALIGNMENTS

##### RESULT 1

A34835  
ribosomal protein L34 - Pseudomonas aeruginosa (fragment)  
C;Species: Pseudomonas aeruginosa  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 23-Jun-1993  
C;Accession: A34835  
R;Yee, T.W.; Smith, D.W.  
Proc. Natl. Acad. Sci. U.S.A. 87, 1278-1282, 1990  
A;Title: Pseudomonas chromosomal replication origins: a bacterial class distinct from  
A;Reference number: A34835; MUID:90160310  
A;Accession: A34835  
A;Molecule type: DNA  
A;Residues: 1-17 <YEE>  
A;Cross-references: GB:M30125  
C;Genetics:  
A;Gene: rpmH  
C;Superfamily: Escherichia coli ribosomal protein L34  
C;Keywords: protein biosynthesis; ribosome

Query Match 26.8%; Score 33; DB 2; Length 17;  
Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 KNGRYSISRTEA 14  
||||: |||  
Db 6 KNGRQVLSRRRA 17

##### RESULT 2

B39845  
pyrB leader peptide - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 12-Dec-1997  
C;Accession: B39845  
R;Quinn, C.L.; Stephenson, B.T.; Switzer, R.L.  
J. Biol. Chem. 266, 9113-9127, 1991  
A;Title: Functional organization and nucleotide sequence of the Bacillus subtilis pyr  
A;Reference number: A39845; MUID:91225016  
A;Accession: B39845  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-19 <QUI>  
A;Cross-references: GB:M59757  
C;Superfamily: unassigned leader peptides

Query Match 22.8%; Score 28; DB 2; Length 19;  
Best Local Similarity 66.7%; Pred. No. 9.6e+02;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 17 CCKAFN 22

Db 8 CCPLFN 13

## RESULT 3

I54351

gene HEXA protein - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jul-2000

C:Accession: I54351

R:Akli, S.; Chomel, J.C.; Lacorte, J.M.; Bachner, L.; Poenaru, L.

Hum. Mol. Genet. 2, 61-67, 1993

A:Title: Ten novel mutations in the HEXA gene in non-Jewish Tay-Sachs patients.

A:Reference number: I54351; MUID:93258352

A:Accession: I54351

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-21 <RES>

A:Cross-references: GB:S61298; NID:g300412; PIDN:AAD13927.1; PID:g4261627

C:Genetics:

A:Gene: GDB:HEXA

A:Cross-references: GDB:L20040; OMIM:272800

A:Map position: 15q23-15q24

C:Superfamily: beta-hexosaminidase

## Query Match

Best Local Similarity 22.0%; Score 27; DB 2; Length 21;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 GRYSISRTAAD 16

I : I : I : I : I

I : I : I : I : I

## RESULT 4

S65435

manganese peroxidase (EC 1.11.1.13) MnP1 - Pleurotus eryngii (fragment)

C:Species: Pleurotus eryngii

C>Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997

C:Accession: S65435

R:Martinez, M.J.; Ruiz-Duenas, F.J.; Guillen, F.; Martinez, A.T.

Eur. J. Biochem. 237, 424-432, 1996

A:Title: Purification and catalytic properties of two manganese peroxidase isoenzymes fr

A:Reference number: S65434; MUID:96215438

A:Accession: S65435

A:Molecule type: protein

A:Residues: 1-19 <MAR>

A:Experimental source: CBS 613.91

C:Keywords: manganese; oxidoreductase

## Query Match

Best Local Similarity 21.5%; Score 26.5; DB 2; Length 19;

Matches 7; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 11 RTEA-ADCCAFN 22

I I I I I I

I I I I I I

## RESULT 5

T06383

hypothetical protein - soybean

C:Species: Glycine max (soybean)

C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 11-May-2000

C:Accession: T06383

R:Dewey, R.E.; Wilson, R.F.; Novitzky, W.P.; Goode, J.H.

Plant Cell 6, 1495-1507, 1994

A:Title: The AAPT1 gene of soybean complements a cholinephosphotransferase-deficient mut

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

A:Accession: T06383

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Keywords: manganese; oxidoreductase

A:Reference number: T06383; MUID:95086383

R;Martinez, M.J.; Ruiz-Duenas, F.J.; Guillen, F.; Martinez, A.T.  
 Eur. J. Biochem. 237, 424-432, 1996  
 A:Title: Purification and catalytic properties of two manganese peroxidase isoenzymes from  
 A:Reference number: S65434; MUID:96215438  
 A:Accession: S65434  
 A:Molecule type: protein  
 A:Residues: 1-19 <MAR>  
 A:Experimental source: CBS 613.91  
 C:Keywords: manganese; oxidoreductase

Query Match 19.9%; Score 24.5; DB 2; Length 19;  
 Best Local Similarity 58.3%; Pred. No. 3.1e+03;  
 Matches 7; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 11 RTEA-ADCKAF 21  
 ||| ||| |||  
 Db 7 RTTADAACCLF 18

RESULT 9  
 S66222  
 A:Title: defens AMP2 - Dahlia merckii (fragment)  
 N:Alternate names: seed antifungal protein  
 C:Species: Dahlia merckii  
 C:Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 17-Mar-1999  
 C:Accession: S66222  
 R;Osborn, R.W.; de Samblanx, G.W.; Thevissen, K.; Goderis, I.; Torrekens, S.; van Leuven  
 FEBS Lett. 368, 257-262, 1995  
 A:Title: Isolation and characterisation of plant defensins from seeds of Asteraceae, Fab  
 A:Reference number: S66218; MUID:95354848  
 A:Accession: S66222  
 A:Molecule type: protein  
 A:Residues: 1-20 <OSB>  
 C:Keywords: antifungal

Query Match 19.9%; Score 24.5; DB 2; Length 20;  
 Best Local Similarity 35.3%; Pred. No. 3.2e+03;  
 Matches 6; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 1 CEKNGRYSISRTEAADC 17  
 ||| ||| |||  
 Db 3 CEK-----ASKTWSGNC 14

RESULT 10  
 D48394  
 A:Title: major fat-globule membrane protein GP 55 - guinea pig (fragment)  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Oct-1997  
 C:Accession: D48394  
 R;Mathier, I.H.; Banghart, L.R.; Lane, W.S.  
 Biochem. Mol. Biol. Int. 29, 545-554, 1993  
 A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
 A:Note: sequence extracted from NCBI backbone (NCBIP:131448)  
 C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo

Query Match 19.5%; Score 24; DB 2; Length 15;  
 Best Local Similarity 80.0%; Pred. No. 3e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEKNG 5  
 ||| |||  
 Db 6 CELNG 10

RESULT 11  
 PQ0544  
 A:Title: capsid protein VP5 - human herpesvirus 1 (fragments)  
 C:Species: human herpesvirus 1  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
 C:Accession: PQ0544  
 R;Davison, M.D.; Rixon, F.J.; Davison, A.J.  
 J. Gen. Virol. 73, 2709-2713, 1992

A:Title: Identification of genes encoding two capsid proteins (VP24 and VP26) of herp  
 A:Reference number: PQ0544; MUID:93019027  
 A:Accession: PQ0544  
 A:Molecule type: protein  
 A:Residues: 1-20 <DAV>  
 A:Experimental source: strain 17  
 C:Genetics:  
 A:Gene: UL19  
 C:Keywords: capsid protein

Query Match 19.5%; Score 24; DB 2; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 NGRYSISRTEAA 15  
 ||| : |||  
 Db 6 NGRILAAPATIAA 17

RESULT 12  
 S67990  
 A:Title: neurotoxin-associated protein type A Hn+ 17K chain - Clostridium botulinum (fragment)  
 N:Alternate names: HA-15 protein  
 C:Species: Clostridium botulinum  
 C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 12-Jun-1998  
 C:Accession: S67990; C44644  
 R;Fujita, K.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.  
 FEBS Lett. 376, 41-44, 1995  
 A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone  
 A:Reference number: S67988; MUID:96096783  
 A:Accession: S67990  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <FOJ>  
 R;Somers, E.; DasGupta, B.R.  
 J. Protein Chem. 10, 415-425, 1991  
 A:Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without  
 A:Reference number: A4644; MUID:92143938  
 A:Contents: type A  
 A:Accession: C44644  
 A:Molecule type: protein  
 A:Residues: 1-19 <SOM>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:83795)  
 C:Keywords: hemagglutinin

Query Match 19.5%; Score 24; DB 2; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 NGRYSI 9  
 ||| : |||  
 Db 9 NGNYNI 14

RESULT 13  
 A56894  
 A:Title: intracrystalline chromoprotein 1 - Terebratella sanguinea (fragment)  
 C:Species: Terebratella sanguinea  
 C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 20-Jun-2000  
 C:Accession: A56894

R;Cusack, M.; Curry, G.; Clegg, H.; Abbott, G.  
Comp. Biochem. Physiol. B 102, 93-95, 1992  
A:Title: An intracrySTALLINE chromatoprotein from red brachiopod shells: implications for  
A:Reference number: A56894; MUID:92405551  
A:Contents: Leach, red brachiopod shells  
A:Accession: A56894  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <CUS>  
A:Note: sequence extracted from NCBI backbone (NCBIP:114882)  
C:Keywords: chromatoprotein

Query Match 19.5%; Score 24; DB 2; Length 20;  
Best Local Similarity 62.5%; Pred. No. 3.8e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 ISRTEAD 16  
|||  
Db 11 ISKTSQAD 18

RESULT 14  
MXKN1  
mu-conotoxin GIIIA [validated] - cone shell (Conus geographus)  
N:Alternate names: geographutoxin I (GTX I); myotoxin I  
C:Species: Conus geographus (geography cone)  
C:Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 15-Sep-2000  
C:Accession: A01786; A23579  
R:Sato, S.; Nakamura, H.; Ohizumi, Y.; Kobayashi, J.; Hirata, Y.  
FEBS Lett. 155, 277-280, 1983  
A:Title: The amino acid sequences of homologous hydroxyproline-containing myotoxins from  
A:Reference number: A91309; MUID:83210170  
A:Accession: A01786  
A:Molecule type: protein  
A:Residues: 1-22 <SAT>  
R:Cruz, L.J.; Gray, W.R.; Olivera, B.M.; Zeikus, R.D.; Kerr, L.; Yoshikami, D.; Moczylo  
J. Biol. Chem. 260, 9280-9288, 1985  
A:Title: Conus geographus toxins that discriminate between neuronal and muscle sodium ch  
A:Reference number: A23579; MUID:85261316  
A:Accession: A23579  
A:Molecule type: protein  
A:Residues: 1-22 <CRU>  
R:Kohda, D.; Lancelin, J.M.; Inagaki, F.; Wakamatsu, K.  
submitted to the Brookhaven Protein Data Bank, December 1992  
A:Reference number: A51994; PDB:1TCG  
A:Contents: annotation; conformation by (1)H-NMR, residues 1-22  
R:Lancelin, J.M.; Kohda, D.; Tate, S.I.; Yanagawa, Y.; Abe, T.; Satake, M.; Inagaki, F.  
Biochemistry 30, 6908-6916, 1991  
A:Title: Tertiary structure of conotoxin GIIIA in aqueous solution.  
A:Reference number: A44659; MUID:91299744  
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
R:Ott, K.H.; Becker, S.; Gordon, R.D.; Rueterjans, H.  
FEBS Lett. 278, 160-166, 1991  
A:Title: Solution structure of mu-conotoxin GIIIA analysed by 2D-NMR and distance geomet  
A:Reference number: A58581; MUID:91122275  
A:Contents: annotation; conformation by (1)H-NMR  
R:Wakamatsu, K.; Kohda, D.; Hatanaka, H.; Lancelin, J.M.; Ishida, Y.; Oya, M.; Nakamura,  
Biochemistry 31, 12577-12584, 1992  
A:Title: Structure-activity relationships of mu-conotoxin GIIIA: structure determination  
A:Reference number: A44244; MUID:93112598  
A:Contents: annotation; conformation by (1)H-NMR  
C:Superfamily: mu-conotoxin  
C:Keywords: amidated carboxyl end; hydroxyproline; myotoxin; sodium channel inhibitor; v  
F:3-15,4-20,10-21/Disulfide bonds: #status experimental  
F:6,7/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental  
F:17/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:22/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 19.5%; Score 24; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DCC 18  
|||  
Db 2 DCC 4

## RESULT 15

MXKN2

mu-conotoxin GIIIB [validated] - cone shell (Conus geographus)

N:Alternate names: geographutoxin II (GTX II); myotoxin II

C:Species: Conus geographus (geography cone)

C:Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 15-Sep-2000

C:Accession: A01787; B23579

R:Sato, S.; Nakamura, H.; Ohizumi, Y.; Kobayashi, J.; Hirata, Y.

FEBS Lett. 155, 277-280, 1983

A:Title: The amino acid sequences of homologous hydroxyproline-containing myotoxins f

A:Reference number: A91309; MUID:83210170

A:Accession: A01787

A:Molecule type: protein

A:Residues: 1-22 &lt;SAT&gt;

R:Cruz, L.J.; Gray, W.R.; Olivera, B.M.; Zeikus, R.D.; Kerr, L.; Yoshikami, D.; Moczy

J. Biol. Chem. 260, 9280-9288, 1985

A:Title: Conus geographus toxins that discriminate between neuronal and muscle sodium

A:Reference number: A23579; MUID:85261316

A:Accession: B23579

A:Molecule type: protein

A:Residues: 1-22 &lt;CRU&gt;

R:Hill, J.M.; Alewood, P.F.; Craik, D.J.

submitted to the Brookhaven Protein Data Bank, April 1996

A:Reference number: A65705; PDB:1GIB

A:Contents: annotation; conformation by (1)H-NMR, residues 1-22

R:Hill, J.M.; Alewood, P.F.; Craik, D.J.

Biochemistry 35, 8824-8835, 1996

A:Title: Three-dimensional solution structure of mu-conotoxin GIIIB, a specific block

A:Reference number: A58590; MUID:96280640

A:Contents: annotation; conformation by (1)H-NMR

C:Superfamily: mu-conotoxin

C:Keywords: amidated carboxyl end; hydroxyproline; myotoxin; sodium channel inhibitor

F:3-15,4-20,10-21/Disulfide bonds: #status experimental

F:6,7/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F:17/Modified site: 4-hydroxyproline (Pro) #status experimental

F:22/Modified site: amidated carboxyl end (Ala) #status experimental

## Query Match

Best Local Similarity 19.5%; Score 24; DB 1; Length 22;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DCC 18

|||

Db 2 DCC 4

Search completed: February 21, 2002, 16:27:20

Job time: 66 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:27:04 ; Search time 10.04 Seconds  
(without alignments)  
80.341 Million cell updates/sec

Title: US-08-753-851-1  
Perfect score: 123  
Sequence: 1 CEKNGRYSISRTEADCKAFN 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1218

Minimum DB seq length: 0  
Maximum DB seq length: 22

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	20.3	22	1 LANM_STRMU	P80666 streptococ
2	24	19.5	22	1 CXM1_CONGE	P01523 conus geogr
3	24	19.5	22	1 CXM2_CONGE	P01524 conus geogr
4	24	19.5	22	1 CXM3_CONGE	P05482 conus geogr
5	22	17.9	15	1 C1QA_RAT	P31720 rattus norv
6	22	17.9	22	1 23KD_BACST	P80166 bacillus st
7	21	17.1	13	1 CXET_CONTE	P81755 conus texti
8	21	17.1	14	1 NEJ2_FASHE	P80526 fasciola he
9	20	16.3	13	1 CXA2_CONGE	P01520 conus geogr
10	20	16.3	14	1 TKNM_RANMA	P40951 rana margar
11	20	16.3	15	1 CXA1_CONGE	P01519 conus geogr
12	20	16.3	19	1 FIBB_HORSE	P14471 equus cabal
13	20	16.3	19	1 HMD_METWO	P32441 methanobact
14	20	16.3	20	1 ALAT_PIG	P13191 sus scrofa
15	20	16.3	21	1 XYNA_DICB4	P80718 dictyoglomu
16	19	15.4	8	1 AL17_CARMA	P81820 carcinus ma
17	19	15.4	9	1 CCAP_CARMA	P38556 carcinus ma
18	19	15.4	12	1 UR2A_CATCO	P04558 catostomus
19	19	15.4	12	1 UR2_GILMI	P01147 gallichthys
20	19	15.4	11	1 HSTB_ECOLI	P01560 escherichia
21	19	15.4	19	1 DHAB_COMTE	P80704 comamonas t
22	19	15.4	21	1 FER_PYRWO	P81638 pyrococcus
23	18.5	15.0	20	1 LPTR_BACSU	P23053 bacillus su
24	18	14.6	12	1 CXAL_CONIM	P50983 conus imper
25	18	14.6	13	1 CXAL_CONST	P15471 conus stria
26	18	14.6	13	1 CXAA_CONST	P28878 conus stria
27	18	14.6	14	1 CXAL_CONCN	P56973 conus conso
28	18	14.6	14	1 CXAL_CONMA	P01521 conus maqus
29	18	14.6	14	1 SMS1_MYOSC	P20750 myoxocephal
30	18	14.6	14	1 SMS_ALLMI	P31885 alligator m
31	18	14.6	15	1 CXA2_CONAL	P56640 conus aulic
32	18	14.6	16	1 CXA1_CONAL	P56639 conus aulic
33	18	14.6	16	1 CXA1_CONEP	P56638 conus episc

## ALIGNMENTS

## RESULT 1

```

LANM_STRMU
ID LANM_STRMU STANDARD; PRT; 22 AA.
AC P80666;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LANTIBIOTIC MUTACIN B-NY266.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE.
RC STRAIN=NY266;
RX MEDLINE=97379322; PubMed=9237644;
RA Motla-Meira M., Lacroix C., Lapointe G., Lavoie M.C.;
RT "Purification and structure of mutacin B-Ny266: a new lantibiotic
   produced by Streptococcus mutans.";
RL FEBS Lett. 410:275-279(1997).
CC -!- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
   ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
   LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
   CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
   TRANSMEMBRANE PORES.
CC -!- MASS SPECTROMETRY: MW=2270.29; MW_ERR=0.21; METHOD=ELECTROSPRAY.
CC -!- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
DR InterPro; IPR001049; Gallidermin.
DR Pfam; PF02052; Gallidermin; 1.
DR PRINTS; PR00323; GALLIDERMIN.
KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.
FT MOD_RES 5 5 DHA (2,3-DIDEHYDROALANINE).
FT MOD_RES 8 8 D-ABU (AMINO-BUTYRIC ACID).
FT MOD_RES 14 14 DHB (2,3-DIDEHYDROBUTYRINE).
FT MOD_RES 19 19 DHA (2,3-DIDEHYDROALANINE).
FT THIOETH 3 7 ALA-S-ALA (LANTHIONINE).
FT THIOETH 8 11 ABU-S-ALA (BETA-METHYLLANTHIONINE).
FT THIOETH 16 21 ALA-S-ALA (LANTHIONINE).
FT THIOETH 19 22 DHA-S-ALA (AVI).
SQ SEQUENCE 22 AA; 2425 MW; 961C1480401F92CE CRC64;

```

Query Match 20.3%; Score 25; DB 1; Length 22;  
Best Local Similarity 27.8%; Pred. No. 8+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 1 CEKNGRYSISRTEADCC 18  
| | | : : : |  
DB 11 CARTGSFN-----SYCC 22

## RESULT 2

```

CXMI_CONGE
ID CXMI_CONGE STANDARD; PRT; 22 AA.
AC P01523;
DT 21-JUL-1986 (Rel. 01, Created)

```

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE MU-CONOTOXIN GIIIA (MYOTOXIN I) (GEOGRAPHUTOXIN I) (GTX-I).  
OS Conus geographus (Geography cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=6491;  
RN [1]  
RN SEQUENCE.  
RP MEDLINE=85261316; PubMed=2410412;  
RX Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,  
RA Yoshikami D., Moczydlowski E.;  
RT "Conus geographus toxins that discriminate between neuronal and  
muscle sodium channels.";  
RL J. Biol. Chem. 260:9280-9288(1985).  
RN [2]  
RN SEQUENCE.  
RP MEDLINE=83210170; PubMed=6852238;  
RX Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;  
RA "The amino acid sequences of homologous hydroxyproline-containing  
myotoxins from the marine snail Conus geographus venom.";  
RL FEBS Lett. 155:277-280(1983).  
RN [3]  
RN DISULFIDE BONDS.  
RP MEDLINE=90249506; PubMed=2338142;  
RX Hidaka Y., Sato K., Nakamura H., Kobayashi J., Ohizumi Y.,  
RA Simonishi Y.;  
RT "Disulfide pairings in geographutoxin I, a peptide neurotoxin from  
Conus geographus.";  
RL FEBS Lett. 264:29-32(1990).  
RN [4]  
RN REVIEW.  
RP MEDLINE=89024586; PubMed=3052286;  
RX Gray W.R., Olivera B.M., Cruz L.J.;  
RA "Peptide toxins from venomous Conus snails.";  
RL Annu. Rev. Biochem. 57:665-700(1988).  
RN [5]  
RN STRUCTURE BY NMR.  
RP MEDLINE=91122275; PubMed=1991506;  
RX Olt K.-H., Becker S., Gordon R.D., Rueterjans H.;  
RA "Solution structure of mu-conotoxin GIIIA analysed by 2D-NMR and  
distance geometry calculations.";  
RL FEBS Lett. 278:160-166(1991).  
RN [6]  
RN STRUCTURE BY NMR.  
RP MEDLINE=91299744; PubMed=2069951;  
RX Lancelin J.-M., Kohda D., Tate S.-I., Yanagawa Y., Abe T., Satake M.,  
RA Inagaki F.;  
RT "Tertiary structure of conotoxin GIIIA in aqueous solution.";  
RL Biochemistry 30:6908-6916(1991).  
CC -1- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
CHANNELS.  
CC PIR; A01786; MKKN1.  
DR PIR; A23579; A23579.  
DR PDB; 1TCG; 31-JAN-94.  
DR PDB; 1TCH; 31-JAN-94.  
DR PDB; 1TCJ; 31-JAN-94.  
DR PDB; 1TCK; 31-JAN-94.  
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;  
3D-structure.  
FT DISULFID 3 15  
FT DISULFID 4 20  
FT DISULFID 10 21  
FT MOD\_RES 6 6 HYDROXYLATION.  
FT MOD\_RES 7 7 HYDROXYLATION.  
FT MOD\_RES 17 17 HYDROXYLATION.  
FT MOD\_RES 22 22 AMIDATION.  
FT HELIX 13 16  
FT TURN 19 21  
SQ SEQUENCE 22 AA; 2568 MW; F6CB02ADB359813C CRC64;

Query Match 19.5%; Score 24; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
QY 16 DCC 18  
DB 2 DCC 4  
RESULT 3  
CAM2\_CONGE STANDARD; PRT; 22 AA.  
ID CXM2\_CONGE AC P01524;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE MU-CONOTOXIN GIIIB (MYOTOXIN II) (GEOGRAPHUTOXIN II) (GTX-II).  
OS Conus geographus (Geography cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=6491;  
RN [1]  
RN SEQUENCE.  
RP MEDLINE=85261316; PubMed=2410412;  
RX Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,  
RA Yoshikami D., Moczydlowski E.;  
RT "Conus geographus toxins that discriminate between neuronal and  
muscle sodium channels.";  
RL J. Biol. Chem. 260:9280-9288(1985).  
RN [2]  
RN SEQUENCE.  
RP MEDLINE=83210170; PubMed=6852238;  
RX Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;  
RA "The amino acid sequences of homologous hydroxyproline-containing  
myotoxins from the marine snail Conus geographus venom.";  
RL FEBS Lett. 155:277-280(1983).  
RN [3]  
RN REVIEW.  
RP MEDLINE=89024586; PubMed=3052286;  
RX Gray W.R., Olivera B.M., Cruz L.J.;  
RA "Peptide toxins from venomous Conus snails.";  
RL Annu. Rev. Biochem. 57:665-700(1988).  
RN [4]  
RN STRUCTURE BY NMR.  
RP MEDLINE=96280640; PubMed=8688418;  
RX Hill J.M., Alewood P.F., Craik D.J.;  
RA "Three-dimensional solution structure of mu-conotoxin GIIIB, a  
specific blocker of skeletal muscle sodium channels.";  
RL Biochemistry 35:8824-8835(1996).  
CC -1- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
CHANNELS.  
CC PIR; A01787; MKKN2.  
DR PIR; B23579; B23579.  
DR PDB; 1GIB; 08-NOV-96.  
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;  
3D-structure.  
FT DISULFID 3 15  
FT DISULFID 4 20  
FT DISULFID 10 21  
FT MOD\_RES 6 6 HYDROXYLATION.  
FT MOD\_RES 7 7 HYDROXYLATION.  
FT MOD\_RES 17 17 HYDROXYLATION.  
FT MOD\_RES 22 22 AMIDATION.  
SQ SEQUENCE 22 AA; 2599 MW; F50402BA93199E01 CRC64;

Query Match 19.5%; Score 24; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
QY 16 DCC 18  
DB 2 DCC 4

```

Db 2 DCC 4
RESULT 4
CXH3_CONGE
ID CXH3_CONGE STANDARD; PRT; 22 AA.
AC P05482;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE MU-CONOTOXIN GIIC.
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=85261316; PubMed=2410412;
RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
RA Yoshikami D., Moczydlowski E.;
RT "Conus geographus toxins that discriminate between neuronal and
RT muscle sodium channels.";
RL J. Biol. Chem. 260:9280-9288(1985).
RN [2]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -1- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK
CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM
CC CHANNELS.
DR PIR; C23579; C23579.
DR HSSP; P01524; IGIB.
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom.
FT DISULFID 3 15 BY SIMILARITY.
FT DISULFID 4 20 BY SIMILARITY.
FT DISULFID 10 21 BY SIMILARITY.
FT MOD_RES 6 6 HYDROXYLATION.
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 17 17 HYDROXYLATION.
FT MOD_RES 22 22 AMIDATION.
FT MOD_RES 22 22 AMIDATION.
SQ SEQUENCE 22 AA; 2553 MW; F50402BA92A9813C CRC64;

Query Match 19.5%; Score 24; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 16 DCC 18
III
Db 2 DCC 4

RESULT 5
C1QA_RAT
ID C1QA_RAT STANDARD; PRT; 15 AA.
AC P31720;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE COMPLEMENT C1Q SUBCOMPONENT, A CHAIN (FRAGMENT).
GN C1QA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE.
RX MEDLINE=93218657; PubMed=8464426;
RA Wing M.G., Seilly D.J., Bridgman D.J., Harrison R.A.;
RT "Rapid isolation and biochemical characterization of rat C1 and C1q.";
RL Mol. Immunol. 30:433-440(1993).

-1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. THE C1Q SUBCOMPONENT IS
CC COMPOSED OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED
CC DIMERS OF THE A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-
CC LINKED DIMERS OF THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C
CC DIMER BANDS, RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
DR InterPro; IPR001073; C1q.
DR PROSITE; PS01113; C1Q; PARTIAL.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat.
FT DISULFID 4 4 INTERCHAIN (WITH CYS-4 IN B CHAIN).
FT NON_TER 15 15
FT SEQUENCE 15 AA; 1488 MW; 1B3D8000B7793965 CRC64;

Query Match 17.9%; Score 22; DB 1; Length 15;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 16 DCCAFN 22
I I I I
Db 2 DVCRAPN 8
I I I I

RESULT 6
23KD_BAGST
ID 23KD_BAGST STANDARD; PRT; 22 AA.
AC P80166;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 23 KDA BASIC PROTEIN (FRAGMENT).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE.
RA Vergias C.E.;
RL Submitted (OCT-1992) to the SWISS-PROT data bank.
FT NON_TER 22 22
FT SEQUENCE 22 AA; 2376 MW; 96C604E42CE0BF6C CRC64;

Query Match 17.9%; Score 22; DB 1; Length 22;
Best Local Similarity 21.4%; Pred. No. 2.3e+03;
Matches 3; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 9 ISRTEAADCCAFN 22
I I I I
Db 9 VSKVDLSEVANAIN 22
I I I I

RESULT 7
CXET_CONTE
ID CXET_CONTE STANDARD; PRT; 13 AA.
AC P81755;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EPSILON-CONOTOXIN TXIX.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]

```

```

RP SEQUENCE, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY NMR.
RC TISSUE=Venom;
RX MEDLINE=99254114; PubMed=10318957;
RA Rigby A.C., Lucas-Meunier E., Kalume D.E., Czerwicz E., Hambe B.,
RA Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,
RA Furie B.C., Furie B., Stenflo J.P.;
RT "A conotoxin from Conus textile with unusual posttranslational
RT modifications reduces presynaptic Ca2+ influx.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).
CC -!- FUNCTION: CONOTOXIN WHICH ACTS AT PRESYNAPTIC MEMBRANES, BLOCKING
CC THE CALCIUM CHANNELS.
CC -!- PTM: O-GLYCAN CONSISTS OF THE DISACCHARIDE GAL-GALNAc.
DR PDB: 1WCT; 08-JUN-99.
KW Presynaptic neurotoxin; Calcium channel inhibitor; Venom; Vitamin K;
KW Gamma-carboxyglutamic acid; Glycoprotein; Hydroxylation; Bromination;
KW 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 9
FT MOD_RES 1 1 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 4 4 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 7 7 BROMINATION.
FT MOD_RES 13 13 HYDROXYLATION.
FT CARBOHYD 10 10 O-LINKED (GALNAc...).
SQ SEQUENCE 13 AA; 1388 MW; 386C9E1C74AFA378 CRC64;

Query Match 17.1%; Score 21; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 16 DCCK 19
Db 1 ECCE 4

RESULT 8
NEJ2_FASHE STANDARD; PRT; 14 AA.
AC P80526;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE NEWLY EXCYSTED JUVENILE PROTEIN 2 (FRAGMENT).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
OC Rhabditophora; Eulcithophora; Revertospermata; Mediofusata;
OC Neodermata; Trematoda; Digenea; Echinostomida; Echinostomata;
OC Fascioloidae; Fasciolidae; Fasciola.
OX NCBI_TaxID=6192;
RN [1]
RP SEQUENCE.
RX MEDLINE=95366993; PubMed=7639732;
RA Tkalcovic J., Ashman K., Meeusen E.;
RT "Fasciola hepatica: rapid identification of newly excysted juvenile
RT proteins.";
RL Biochem. Biophys. Res. Commun. 213:169-174(1995).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT THE NEWLY EXCYSTED JUVENILE
CC STAGE.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1581 MW; 9E0F0090CC8C0DF1 CRC64;

Query Match 17.1%; Score 21; DB 1; Length 14;
Best Local Similarity 80.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EKNGR 6
Db 2 EDNGR 6

RESULT 9
CXA2_CONGE
```

```

ID CXA2_CONGE STANDARD; PRT; 13 AA.
AC P01520;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE ALPHA-CONOTOXIN GII.
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=81191854; PubMed=7014556;
RA Gray W.R., Lucue A., Olivera B.M., Barrett J., Cruz L.J.;
RT "Peptide toxins from Conus geographus venom.";
RL J. Biol. Chem. 256:4734-4740(1981).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin MI. Disulfide bonding and conformational states.";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PIR: A01783; NTKN2G.
DR HSSP; P01519; 1NOT.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13 AMIDATION.
SQ SEQUENCE 13 AA; 1422 MW; DEE831C39297EBD CRC64;

Query Match 16.3%; Score 20; DB 1; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.8e+03;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 DCC 18
Db 1 ECC 3

RESULT 10
TKNM_RANNA STANDARD; PRT; 14 AA.
ID TKNM_RANNA
AC P40951;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RANAMARGARIN.
OS Rana margaratae (Chinese frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=121156;
RN [1]
RP SEQUENCE.
RX MEDLINE=90026852; PubMed=2803524;
RA Tang Y.Q., Tian S.H., Wu S.X., Hua J.C., Wu G.F., Zhao E.M., Lu Y.A.,
RA Zhu Y.Q., Zou G., Tsou K.;
RT "Isolation and structure of ranamargarin, a new tachykinin from the
RT skin of Chinese frog Rana margaratae.";
RL Sci. China, B, Chem. Life Sci. 32:570-579(1989).
RN [2]
RP SYNTHESIS.
```

```

RX MEDLINE=90253600; PubMed=2340087;
RA Lu Y.A., Peng J.L., Zhu Y.O., Wu S.X., Tang Y.O., Tian S.H., Zou G.;
RT "Synthesis and biological activity of a new frog skin peptide,
RT ranamarginin.";
RL Sci. China, B, Chem. Life Sci. 33:170-177(1990).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR InterPro; IPR002040; Tachykinin.
DR PROSITE; PS00267; TACHYKININ; 1.
FW Tachykinin; Neuropeptide; Amidation; Amphibian skin.
KT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1617 MW; D4593AE408C3673D CRC64;

Query Match 16.3%; Score 20; DB 1; Length 14;
Best Local Similarity 44.4%; Pred. No. 3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 13 EAADCCKAF 21
DB 2 DASDRAKF 10

RESULT 11
ID CXAL_CONGE STANDARD; PRT; 15 AA.
AC P01519;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-CONOTOXIN GIA [CONTAINS: ALPHA-CONOTOXIN GI (G1)].
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=81191854; PubMed=7014556;
RA Gray W.R., Luque A., Olivera B.M., Barrett J., Cruz L.J.;
RT "Peptide toxins from Conus geographus venom.";
RL J. Biol. Chem. 256:4734-4740(1981).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin M1. Disulfide bonding and conformational states.";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
RL Annu. Rev. Biochem. 57:665-700(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS) OF GI.
RX MEDLINE=96378624; PubMed=8784187;
RA Guddat L.W., Martin J.A., Shan L., Edmondson A.B., Gray W.R.;
RT "Three-dimensional structure of the alpha-conotoxin GI at 1.2-A
RT resolution.";
RL Biochemistry 35:11329-11335(1996).
RN [5]
RP STRUCTURE BY NMR OF GI.
RX MEDLINE=89375269; PubMed=2775719;
RA Pardi A., Galdes A., Florance J., Maniconte D.;
RT "Solution structures of alpha-conotoxin GI determined by two-
RT dimensional NMR spectroscopy.";
RL Biochemistry 28:5494-5501(1989).
RN [6]
RP STRUCTURE BY NMR OF GI.
RX MEDLINE=98239743; PubMed=9571060;

```

```

RA Gehrman J., Alewood P.F., Craik D.J.;
RT "Structure determination of the three disulfide bond isomers of
RT alpha-conotoxin GI: a model for the role of disulfide bonds in
RT structural stability.";
RL J. Mol. Biol. 278:401-415(1998).
RN [7]
RP STRUCTURE BY NMR OF AN ANTITOXIC ANALOG OF GI.
RX MEDLINE=99438341; PubMed=10508392;
RA Mok K.H., Han K.H.;
RT "NMR solution conformation of an antitoxic analogue of alpha-conotoxin
RT GI: identification of a common nicotinic acetylcholine receptor
RT alpha(1)-subunit binding surface for small ligands and alpha-
RT conotoxins.";
RL Biochemistry 38:11895-11904(1999).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC -!- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF CONOTOXIN GIA.
DR PIR; A01782; NTKNAG.
DR PDB; 1NOT; 07-DEC-96.
DR PDB; 1XGA; 16-FEB-99.
DR PDB; 1XGB; 16-FEB-99.
DR PDB; 1XGC; 23-MAR-99.
DR PDB; 1Q53; 06-OCT-99.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; 3D-structure.
FT PEPTIDE 1 15 ALPHA-CONOTOXIN GIA.
FT PEPTIDE 1 13 ALPHA-CONOTOXIN GI.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13 AMIDATION (IN GI).
FT MOD_RES 15 15 AMIDATION (IN GIA).
SQ SEQUENCE 15 AA; 1628 MW; 2AE73EE90F8C2E19 CRC64;

Query Match 16.3%; Score 20; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 3.2e+03;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 DCC 18
DB 1 ECC 3

RESULT 12
FIBB_HORSE
ID FIBB_HORSE STANDARD; PRT; 19 AA.
AC F14471;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last annotation update)
DE FIBRINOPEPTIDE B.
OS Equus caballus (Horse), and Equus asinus (Donkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796, 9793;
RN [1]
RP SEQUENCE.
RC SPECIES=Horse;
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN [2]
RP SEQUENCE.
RC SPECIES=E.asinus;
RA Blomback B., Blomback M., Grondahl N.J., Holmberg E.;
RT "Structure of fibrinopeptides-its relation to enzyme specificity and
RT phylogeny and classification of species.";
RL Ark. Kemi 25:411-428(1966).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

```

```

CC      (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC      -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC      THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC      CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC      RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR      InterPro: IPR002181; Fibrinogen.C
DR      PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW      Blood coagulation; Plasma; Sulfation.
FT      MOD_RES       3      19
FT      NON_TER       3      19
SQ      SEQUENCE      19 AA; 2296 MW; 921A2B02D5F6891D CRC64;

Query Match      16.3%; Score 20; DB 1; Length 19;
Best Local Similarity 33.3%; Pred. No. 4e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 EKNCRYSIS 10
Db      7 EEDGRTKVT 15

RESULT 13
HMD_METWO
ID      HMD_METWO      STANDARD;      PRT;      19 AA.
AC      F32441;
DT      01-OCT-1993 (Rel. 27, Created)
DT      01-OCT-1993 (Rel. 27, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      H(2)-FORMING N5,N10-METHYLENETETRAHYDROMETHANOPTERIN DEHYDROGENASE
DE      (EC 1.5.99.11) (H2-DEPENDENT METHYLENE-H4MPT DEHYDROGENASE)
DE      (FRAGMENT).
OS      HMD.
GN      Methanobacterium wolfei.
OC      Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC      Methanothermobacter.
OX      NCBI_TaxID=145261;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=92394151; PubMed=1521540;
RA      Zhirgibl C., van Dongen W., Schwoerer B., von Buena R.,
RA      Richter M., Klein A., Thauer R.K.;
RT      *H2-forming methylenetetrahydromethanopterin dehydrogenase, a novel
RT      type of hydrogenase without iron-sulfur clusters in methanogenic
RT      archaea";
RL      Eur. J. Biochem. 208:511-520(1992).
CC      -!- CATALYTIC ACTIVITY: N(5),N(10)-METHYLENETETRAHYDROMETHANOPTERIN +
CC      REDUCED COENZYME F(420) = 5-METHYL-5,6,7,8-TETRAHYDROMETHANOPTERIN
CC      + COENZYME F(420).
CC      -!- COFACTOR: ZINC (BY SIMILARITY).
CC      -!- PATHWAY: INVOLVED IN METHANOGENESIS.
CC      -!- SUBUNIT: HOMODIMER.
KW      Oxidoreductase; Methanogenesis; Zinc.
FT      NON_TER       1      1
FT      SEQUENCE      19 AA; 1911 MW; 0C17E9D7BF1F97C9 CRC64;

Query Match      16.3%; Score 20; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 4e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      11 RTEAA 15
Db      12 RTDAA 16

RESULT 14
ALAT_PIG
ID      ALAT_PIG      STANDARD;      PRT;      20 AA.
AC      F13191;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)

```

```

DE      ALANINE AMINOTRANSFERASE (EC 2.6.1.2) (GLUTAMIC--PYRUVIC TRANSAMINASE)
DE      (GPT) (GLUTAMIC--ALANINE TRANSAMINASE) (FRAGMENT).
GN      GPT OR GPT1 OR AAT1.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      SEQUENCE.
RX      TISSUE=Heart;
RX      MEDLINE=79232426; PubMed=465450;
RA      Tanase S., Kojima H., Morino Y.;
RT      "Pyridoxal 5'-phosphate binding site of pig heart alanine
RT      aminotransferase.";
RL      Biochemistry 18:3002-3007(1979).
CC      -!- FUNCTION: PARTICIPATES IN CELLULAR NITROGEN METABOLISM AND ALSO
CC      IN LIVER GLUCONEOGENESIS STARTING WITH PRECURSORS TRANSPORTED
CC      FROM SKELETAL MUSCLES.
CC      -!- CATALYTIC ACTIVITY: L-ALANINE + 2-OXOGLUTARATE = PYRUVATE +
CC      L-GLUTAMATE.
CC      -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC      -!- SUBUNIT: HOMODIMER.
CC      -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC      -!- MISCELLANEOUS: THERE ARE TWO ISOZYMES: A CYTOPLASMIC ONE AND A
CC      MITOCHONDRIAL ONE.
CC      -!- SIMILARITY: TO OTHER SPECIES ALANINE AMINOTRANSFERASE.
DR      PIR: A14344; A14344.
KW      Transferase; Aminotransferase; Pyridoxal phosphate.
FT      NON_TER       1      1
FT      BINDING      11      11 PYRIDOXAL PHOSPHATE.
FT      NON_TER      20      20
FT      SEQUENCE      20 AA; 2217 MW; 1C2243A373EC4801 CRC64;

Query Match      16.3%; Score 20; DB 1; Length 20;
Best Local Similarity 27.3%; Pred. No. 4.2e+03;
Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      7 YSISRTAADC 17
Db      7 HVSXGFGMGC 17

RESULT 15
XYNA_DICB4
ID      XYNA_DICB4      STANDARD;      PRT;      21 AA.
AC      P80718;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      ENDO-1,4-BETA-XYLANASE A (EC 3.2.1.8) (XYLANASE A)
DE      (1,4-BETA-D-XYLAN XYLANOXYLASE A) (FRAGMENT).
OS      Dictyoglomus sp. (strain B4).
OC      Bacteria; Dictyoglomus group; Dictyoglomus.
OX      NCBI_TaxID=69007;
RN      [1]
RP      SEQUENCE.
RA      Adamsen A.K., Jacobsen S., Ahning B.K.;
RA      Submitted (OCT-1996) to the SWISS-PROT data bank.
CC      -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC      LINKAGES IN XYLANS.
CC      -!- PATHWAY: XLAN DEGRADATION.
CC      -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC      HYDROLASES).
DR      InterPro: IPR01000; Glyco_hydro_10.
DR      PROSITE: PS00591; GLYCOSYL_HYDROL_F10; PARTIAL.
KW      Xylan degradation; Hydrolase; Glycosidase.
FT      NON_TER      21      21
FT      SEQUENCE      21 AA; 2406 MW; D4D8E194B5707CA9 CRC64;

Query Match      16.3%; Score 20; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 4.4e+03;

```

1

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NGRYSI 9

Db 3 NERFSI 8

Search completed: February 21, 2002, 16:30:04  
Job time: 180 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:26:44 ; Search time 22.08 Seconds  
(without alignments)  
145.742 Million cell updates/sec

Title: US-08-753-851-1  
Sequence: 1 CERNGRYSISRTAADCCKAFN 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 6292

Minimum DB seq length: 0  
Maximum DB seq length: 22

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL.17.\*
- 1: sp-archaea.\*
  - 2: sp-bacteria.\*
  - 3: sp-fungi.\*
  - 4: sp-human.\*
  - 5: sp-invertebrate.\*
  - 6: sp-mammal.\*
  - 7: sp-mhc.\*
  - 8: sp-organelle.\*
  - 9: sp-phage.\*
  - 10: sp-plant.\*
  - 11: sp-rodent.\*
  - 12: sp-virus.\*
  - 13: sp-vertebrate.\*
  - 14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	24.4	18	Q9UGN8	Q9ugn8 homo sapien
2	30	24.4	21	Q9BP52	Q9bp52 conus arena
3	28	22.8	22	Q9N1W5	Q9n1w5 equus cabal
4	27	22.0	18	Q9MY73	Q9my73 homo sapien
5	27	22.0	19	Q9DE21	Q9de21 gallus gall
6	27	22.0	21	Q16017	Q16017 homo sapien
7	26	21.1	12	Q9UMR0	Q9umr0 homo sapien
8	26	21.1	18	Q9BLD8	Q9bl88 lingua ung
9	25	20.3	11	Q9NL65	Q9nl65 ascaris suu
10	25	20.3	16	Q9TR06	Q9tr06 bos taurus
11	25	20.3	17	Q9R581	Q9r581 vibrio chol
12	25	20.3	18	Q9R580	Q9r580 vibrio chol
13	25	20.3	19	Q9R579	Q9r579 vibrio chol
14	25	20.3	22	Q9S897	Q9s897 citrus sine
15	24	19.5	13	Q9S897	Q9s897 citrus sine
16	24	19.5	18	Q9T298	Q9t298 solanum tub
17	24	19.5	19	Q9R5N8	Q9r5n8 clostridium
18	24	19.5	20	Q9UC14	Q9uc14 homo sapien
19	24	19.5	20	Q9TRQ3	Q9trq3 bos taurus

20	24	19.5	21	10	Q41580	Q41580 triticum ae
21	23	18.7	13	2	Q9AIR1	Q9air1 pseudomonas
22	23	18.7	15	8	Q9T2G9	Q9t2g9 solanum tub
23	23	18.7	18	8	Q78379	Q78379 theileria a
24	23	18.7	19	5	Q9TW06	Q9twq6 tachypleus
25	23	18.7	20	8	Q36005	Q36005 trypanosoma
26	23	18.7	20	11	Q9QV61	Q9qv61 rattus sp.
27	23	18.7	21	8	Q35998	Q35998 trypanosoma
28	22	17.9	11	7	Q29831	Q29831 homo sapien
29	22	17.9	15	4	P78533	P78533 homo sapien
30	22	17.9	16	11	Q9QV05	Q9qv05 rattus sp.
31	22	17.9	18	5	Q9TWS7	Q9tws7 lamellibrac
32	22	17.9	18	9	Q38571	Q38571 bacterioph
33	22	17.9	19	12	Q83965	Q83965 influenza a
34	22	17.9	21	2	Q9R4T3	Q9rat3 mycobacteri
35	22	17.9	21	5	Q25087	Q25087 hermania m
36	22	17.9	21	11	O88229	O88229 mus musculu
37	22	17.9	22	4	Q13726	Q13726 homo sapien
38	22	17.9	22	11	Q9QWB6	Q9qw66 mus sp. sgp
39	21.5	17.5	20	12	Q36978	Q36978 human papil
40	21	17.1	11	11	Q99N81	Q99n81 mus musculu
41	21	17.1	15	4	Q9UCB5	Q9uce5 homo sapien
42	21	17.1	15	7	Q9TNP2	Q9tnp2 homo sapien
43	21	17.1	17	2	O53326	O53326 versinia pe
44	21	17.1	18	2	Q9F581	Q9f581 escherichia
45	21	17.1	18	4	Q9UCN1	Q9ucn1 homo sapien

ALIGNMENTS

RESULT 1

Q9UGN8	Q9UGN8	PRELIMINARY;	PRT;	18 AA.
AC	Q9UGN8;			
DT	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
DT	01-MAY-2000 (TREMBlrel. 13, Last annotation update)			
DE	APAF1 PROTEIN (FRAGMENT).			
GN	APAF1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Roberts D.L., Dalglish R., Cohen G.M., MacFarlane M.;			
RT	"The mammalian CED4 homologue, APAF1, exists as two distinct forms in human cells."			
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AJ133645; CAB65087.1;			
FT	NON_TER	1		
FT	NON_TER	18		
SQ	SEQUENCE	18 AA;	2045 MW;	30D5FA30B885AEF5 CRC64;

Query Match 24.4%; Score 30; DB 4; Length 18;  
Best Local Similarity 62.5%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 SRTEADC 17  
| | | | |  
Db 6 SRSKVADC 13

RESULT 2

Q9BP52	Q9BP52	PRELIMINARY;	PRT;	21 AA.
ID	Q9BP52			
AC	Q9BP52;			
DT	01-JUN-2001 (TREMBlrel. 17, Created)			
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)			
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
DE	CONOTOXIN SCAFFOLD IX PRECURSOR (FRAGMENT).			

OS Conus arenatus.  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=89451;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21105969; PubMed=11158371;  
 RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,  
 RA Fainzilber M.;  
 RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides";  
 RL Mol. Biol. Evol. 18:120-131(2001).  
 DR EMBL; AF215101; AAG60522.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 21 AA; 2298 MW; 2F5672AB6B7171E9 CRC64;

Query Match 24.4%; Score 30; DB 5; Length 21;  
 Best Local Similarity 66.7%; Pred. No. 4.6e+02;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 17 CCKAFN 22  
 II II  
 DB 8 CCSTFN 13

RESULT 3

O9N1W5 PRELIMINARY; PRT; 22 AA.

AC O9N1W5;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE DOPAMINE RECEPTOR D2 (FRAGMENT).  
 GN DRD2.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20082971; PubMed=10613847;  
 RA Caetano A.R., Shlue Y.L., Lyons L.A., O'Brien S.J., Laughlin T.F.,  
 RA Bowling A.T., Murray J.D.;  
 RT "A comparative gene map of the horse (Equus caballus).";  
 RL Genome Res. 9:1239-1249(1999).  
 DR EMBL; AF134062; AAF31305.1; -;  
 KW Receptor.  
 FT NON\_TER 1  
 FT NON\_TER 22  
 SQ SEQUENCE 22 AA; 2643 MW; 2431653B0535E8EB CRC64;

Query Match 22.8%; Score 28; DB 6; Length 22;  
 Best Local Similarity 54.5%; Pred. No. 1e+03;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CEKNRYSISR 11  
 I I I I I I  
 DB 2 CVYNTRYSKR 12

RESULT 4

O9MY73 PRELIMINARY; PRT; 18 AA.

AC O9MY73;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE HUMAN LEUCOCYTE ANTIGEN B.  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bettens F., Tiercy J.M.;  
 RT "Sequence of a new HLA-B\*44Null allele.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ251593; CAB95661.1; -;  
 KW MHC.  
 SQ SEQUENCE 18 AA; 2021 MW; 50D9ADAD2C326362 CRC64;

Query Match 22.0%; Score 27; DB 7; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 11 RTEAADCC 18  
 I I I I I I  
 DB 5 RPEPSSCC 12

RESULT 5

O9DE21 PRELIMINARY; PRT; 19 AA.

AC O9DE21;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE UORF4.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20528616; PubMed=11073974;  
 RA Kobayashi M., Yu R.T., Yasuda K., Umesono K.;  
 RT "Cell-type-specific regulation of the retinoic acid receptor mediated  
 by the orphan nuclear receptor TLX.";  
 RL Mol. Cell. Biol. 20:8731-8739(2000).  
 DR EMBL; AF220160; AAG35365.1; -;  
 SQ SEQUENCE 19 AA; 2316 MW; 2CB3FB243FD48C28 CRC64;

Query Match 22.0%; Score 27; DB 13; Length 19;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CEKNRGR 6  
 I I I I I I  
 DB 13 CQKRGR 18

RESULT 6

O16017 PRELIMINARY; PRT; 21 AA.

AC O16017;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE HEXA PROTEIN (FRAGMENT).  
 GN HEXA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93258352; PubMed=8490625;  
 RA Akli S., Chomel J.C., Lacorte J.M., Bachner L., Poenaru A.,  
 RA Poenaru L.;  
 RT "Ten novel mutations in the HEXA gene in non-Jewish Tay-Sachs  
 patients.";

RL Hum. Mol. Genet. 2:61-67(1993).  
 DR EMBL; S61298; AAD13927.1; -.  
 DR HSSP; P06865; IQBC.  
 FT NON\_TER 1  
 SQ SEQUENCE 21 AA; 2494 MW; D4ACE2D1DA24D8EC CRC64;

Query Match 22.0%; Score 27; DB 4; Length 21;  
 Best Local Similarity 41.7%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 GYISIRTEAAD 16  
 I : I : I I I  
 Db 1 GTFFINKTEIED 12

RESULT 7  
 Q9UMR0 PRELIMINARY; PRT; 12 AA.  
 AC Q9UMR0;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE FRUCTOSE-1,6-BISPHOSPHATASE (EC 3.1.3.11) (FRAGMENT).  
 GN FBP2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tillmann H., Eschrich K.;  
 RT "Structure of human FBP2 gene."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ238483; CAB53359.1; -.  
 DR HSSP; P06636; 1FRP.  
 KW Hydrolase.  
 FT NON\_TER 1 12  
 SQ SEQUENCE 12 AA; 1382 MW; 4CEB259E57386403 CRC64;

Query Match 21.1%; Score 26; DB 4; Length 12;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EKNGRYSI 9  
 I I I I I  
 Db 4 EKRGYVV 11

RESULT 8  
 Q9BLD8 PRELIMINARY; PRT; 18 AA.  
 AC Q9BLD8;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE ELONGATION FACTOR 1 ALPHA (FRAGMENT).  
 GN EF-1A.  
 OS Lingula unguis.  
 OC Eukaryota; Metazoa; Brachiopoda; Linguliformea; Lingulida;  
 OC Linguloidea; Lingulidae; Lingula.  
 OX NCBI\_TaxID=7574;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A15;  
 RA Endo K., Ozawa T., Kojima S.;  
 RT "Nuclear and mitochondrial gene sequence reveal unexpected genetic  
 RT heterogeneity among northern Pacific populations of the brachiopod  
 RT Lingula anatina."  
 RL Mar. Biol. 0:0-0(2001).  
 DR EMBL; AB056463; BAB33373.1; -.

FT NON\_TER 1 1  
 FT NON\_TER 18 18  
 SQ SEQUENCE 18 AA; 1835 MW; D41519D8611680C8 CRC64;

Query Match 21.1%; Score 26; DB 5; Length 18;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ISRTEAADC 17  
 I : I I I I  
 Db 4 ITGTSQADC 12

RESULT 9  
 Q9NL65 PRELIMINARY; PRT; 11 AA.  
 AC Q9NL65;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE ASABF-DELTA (FRAGMENT).  
 GN ASABF-DELTA.  
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;  
 OC Ascarididae; Ascaris.  
 OX NCBI\_TaxID=6253;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kato Y.;  
 RT "Ascaris suum asabf-delta gene, exon 2";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB029815; BAA89496.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1187 MW; 8BADD0CD1EAB5861 CRC64;

Query Match 20.3%; Score 25; DB 5; Length 11;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEKNG 5  
 I I I I  
 Db 7 CEKRG 11

RESULT 10  
 Q9TR06 PRELIMINARY; PRT; 16 AA.  
 AC Q9TR06;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE NEURON-SPECIFIC ASH/GRB-2 SH3 DOMAIN-BINDING PROTEIN (FRAGMENT).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96257761; PubMed=8687411;  
 RA Miura K., Miki H., Shimazaki K., Kawai N., Takenawa T.;  
 RT "Interaction of Ash/Grb-2 via its SH3 domains with neuron-specific  
 RT p150 and p65."  
 RL Biochem. J. 316:639-645(1996).  
 SQ SEQUENCE 16 AA; 1827 MW; E8B5540CD9C828E CRC64;

Query Match 20.3%; Score 25; DB 6; Length 16;  
 Best Local Similarity 55.6%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 NGRYSISRT 12  
: | | | |  
Db 1 DGARSVSRT 9

RESULT 11  
Q9R581 ID Q9R581 PRELIMINARY; PRT; 17 AA.  
AC Q9R581;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE 01-ST-1, NAG-ST, VM-ST-HEAT-STABLE ENTEROTOXIN.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93314823; PubMed=8325391;  
RA Yoshino K., Miyachi M., Takao T., Bag P.K., Huang X., Nair G.B.,  
RA Takeda T., Shimonishi Y.;  
RT "Purification and sequence determination of heat-stable enterotoxin  
RT elaborated by a cholera toxin-producing strain of Vibrio cholerae  
RT 01.";  
RL FEBS Lett. 326:83-86(1993).  
DR InterPro; IPR001489; Enterotoxin\_HS.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
SQ SEQUENCE 17 AA; 1821 MW; 30FF036D018D601C CRC64;

Query Match 20.3%; Score 25; DB 2; Length 17;  
Best Local Similarity 75.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 DCKE 19  
: | | | |  
Db 2 DCCE 5

RESULT 12  
Q9R580 ID Q9R580 PRELIMINARY; PRT; 18 AA.  
AC Q9R580;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE 01-ST-2, VC-H-ST-HEAT-STABLE ENTEROTOXIN.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93314823; PubMed=8325391;  
RA Yoshino K., Miyachi M., Takao T., Bag P.K., Huang X., Nair G.B.,  
RA Takeda T., Shimonishi Y.;  
RT "Purification and sequence determination of heat-stable enterotoxin  
RT elaborated by a cholera toxin-producing strain of Vibrio cholerae  
RT 01.";  
RL FEBS Lett. 326:83-86(1993).  
DR InterPro; IPR001489; Enterotoxin\_HS.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
SQ SEQUENCE 18 AA; 1934 MW; 3080692D018D601C CRC64;

Query Match 20.3%; Score 25; DB 2; Length 18;  
Best Local Similarity 75.0%; Pred. No. 2.5e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 DCKE 19  
: | | | |  
Db 3 DCCE 6

RESULT 13  
Q9R579 ID Q9R579 PRELIMINARY; PRT; 19 AA.  
AC Q9R579;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE 01-ST-3-HEAT-STABLE ENTEROTOXIN.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93314823; PubMed=8325391;  
RA Yoshino K., Miyachi M., Takao T., Bag P.K., Huang X., Nair G.B.,  
RA Takeda T., Shimonishi Y.;  
RT "Purification and sequence determination of heat-stable enterotoxin  
RT elaborated by a cholera toxin-producing strain of Vibrio cholerae  
RT 01.";  
RL FEBS Lett. 326:83-86(1993).  
DR InterPro; IPR001489; Enterotoxin\_HS.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
SQ SEQUENCE 19 AA; 2048 MW; 308015F1A18D601C CRC64;

Query Match 20.3%; Score 25; DB 2; Length 19;  
Best Local Similarity 75.0%; Pred. No. 2.7e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 DCKE 19  
: | | | |  
Db 4 DCCE 7

RESULT 14  
Q9S897 ID Q9S897 PRELIMINARY; PRT; 22 AA.  
AC Q9S897;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE 2N-BINDING PROTEIN (FRAGMENT).  
OS Citrus sinensis (Sweet orange).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Sapindales; Rutaceae; Citrus.  
OX NCBI\_TaxID=2711;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96351190; PubMed=8742339;  
RA Taylor K.C., Albrigo L.G., Chase C.D.;  
RT "Purification of a Zn-binding phloem protein with sequence identity to  
RT chitin-binding proteins.";  
RL Plant Physiol. 110:657-664(1996).  
DR HSP; P27275; LMWC.  
SQ SEQUENCE 22 AA; 2252 MW; 250CF86C0C32B748 CRC64;

Query Match 20.3%; Score 25; DB 10; Length 22;  
Best Local Similarity 50.0%; Pred. No. 3.1e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 DCKRAF 21  
: | | | |  
Db 12 ECCSRF 17

RESULT 15  
O86260 ID O86260 PRELIMINARY; PRT; 13 AA.  
AC O86260;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE TRXB PROTEIN (FRAGMENT).  
 GN TRXB.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 OX NCBI\_TaxID=573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MSA1;  
 RA Albrecht C., Kleiner D.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y17587; CAA76799.1; -  
 ET NON\_TER 1  
 SQ SEQUENCE 13 AA; 1425 MW; 890C8715ECD2287A CRC64;

Query Match 19.5%; Score 24; DB 2; Length 13;  
 Best Local Similarity 80.0%; Pred No. 2.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 15 ADCK 19  
 Db 9 ADACK 13

Search completed: February 21, 2002, 16:29:48  
 Job time: 184 sec

GenCore version 4.5  
Copyright (C) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:26:09 ; Search time 23.83 Seconds  
(without alignments)  
68.385 Million cell updates/sec

Title: US-08-753-851-1

Perfect score: 123

Sequence: 1 CEKNGRYSISRTAADCCKAFN 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 210642

Minimum DB seq length: 0

Maximum DB seq length: 22

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	100.0	22	15 AAR53473	CD44 peptide CD44-
2	43	35.0	22	21 AAY57162	CamK peptide fragm
3	37	30.1	19	9 AAR80554	Partial sequence o
4	37	30.1	19	10 AAP90941	Residues 35-53 of
5	37	30.1	19	15 AAR49699	Cattle, sheep GH p
6	36	29.3	19	9 AAR80556	Partial sequence o
7	36	29.3	19	10 AAP90943	Variant of residue
8	36	29.3	19	15 AAR49700	Pig GH peptide 35-
9	36	29.3	21	18 AAW35265	Epitope comprising
10	35	28.5	19	20 AAY05067	Tumour antigen ant
11	34	27.6	15	17 AAR96887	Human cartilage li

12	34	27.6	17	22 AAG62706	Hanatoxin-like seq
13	34	27.6	19	18 AAW35266	Epitope comprising
14	33	26.8	13	21 AAY99910	Peptide encoded by
15	33	26.8	20	21 AAY97089	Tau conotoxin sequ
16	33	26.8	20	21 AAY97107	Tau conotoxin sequ
17	33	26.8	22	22 AAG62723	Amino acid sequenc
18	32	26.0	15	17 AAR96892	Human proteoglycan
19	32	26.0	16	19 AAW68318	MHC binding peptid
20	32	26.0	16	21 AAB09114	Hepatitis GB virus
21	32	26.0	16	21 AAY68195	Altered MHC determ
22	32	26.0	16	21 AAY52849	Altered MHC determ
23	32	26.0	16	22 AAB58610	Altered MHC determ
24	32	26.0	19	19 AAW53789	KGF-2 deletion mut
25	32	26.0	19	20 AAY32892	KGF-2 antigenic re
26	32	26.0	19	21 AAB10306	Human KGF-2 antige
27	32	26.0	19	22 AAB61670	KGF-2 antigenic re
28	32	26.0	22	20 AAY36618	Fragment of human
29	32	26.0	22	21 AAY87217	Human secreted pro
30	32	26.0	22	22 AAE06199	Human gene 65 enco
31	31	25.2	7	19 AAW60415	Tumour homing pept
32	31	25.2	7	20 AAW93742	Mouse B16B13b mela
33	31	25.2	7	21 AAB21832	Murine melanoma ho
34	31	25.2	7	22 AAE06410	Tumour homing pept
35	31	25.2	20	14 AAR41487	Peptide B11 derive
36	31	25.2	22	15 AAR53591	Pig somatotropin t
37	30	24.4	15	13 AAR26097	Proposed Superanti
38	30	24.4	15	18 AAW36061	Immunogenic epitop
39	30	24.4	15	20 AAY23788	Binding agent homo
40	30	24.4	15	20 AAY25446	T cell receptor be
41	30	24.4	15	22 AAB68377	Antigen binding si
42	30	24.4	16	22 AAG62708	Hanatoxin-like seq
43	30	24.4	17	22 AAG62703	Hanatoxin-like seq
44	30	24.4	17	22 AAG62707	Hanatoxin-like seq
45	30	24.4	21	20 AAY00357	Fragment of human

#### ALIGNMENTS

RESULT 1

ID AAR53473 standard; peptide; 22 AA.

XX AAR53473;

XX AC

DT 01-DEC-1994 (first entry)

XX

DE CD44 peptide CD44-1.

XX

XX Cell adhesion molecule; CD44; antinflammatory; rheumatoid;

KW arthritis; tumor cell metastasis; autoimmune disease;

KW immunosuppressive.

XX

OS Homo sapiens.

XX

PN WO9409811-A.

XX

PD 11-MAY-1994.

XX

PF 29-OCT-1993; 93WO-US10412.

XX

PR 30-OCT-1992; 92US-0973339.

XX

PA (UYDU-) UNIV DUKE.

XX

PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;

XX

DR WPI; 1994-167121/20.

XX

PT Use of CD44 protein and new peptide derivs - for developing prods

PT for inflammation, immune-mediated tissue damage and tumour cell

PT metastasis

XX

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 37-57 of the CD44 protein sequence.

XX Sequence 22 AA;

Query Match 100.0%; Score 123; DB 15; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.7e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEKNGRYSISRTEADCKAEN 22  
Db 1 cekngrysisrteadckafn 22

RESULT 2

AAAY57162  
ID AAY57162 standard; peptide; 22 AA.

XX AC AAY57162;

DT 11-FEB-2000 (first entry)

XX CAMK peptide fragment (residues 281-302).

DE Calmodulin-dependent phosphoenzyme II; CamKII; inhibitor.

XX Synthetic.

PN JP11152298-A.

PD 08-JUN-1999.

PF 19-NOV-1997; 97JP-0336351.

PR 19-NOV-1997; 97JP-0336351.

XX (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.

XX WPI; 2000-026565/03.

XX A specific peptide inhibitor for - calmodulin-dependent phosphoenzyme II

PS Disclosure; Page 2; 5pp; Japanese.

XX The invention provides a peptide having an amino acid sequence KXLRQAEAFDAY where X = Ala or Lys. The peptide is an excellent inhibitor for calmodulin-dependent phosphoenzyme II (CamKII). Sequences AAY57160-161 are CamKII inhibitor peptides and represent specific examples of the generic peptide given above. The present sequence represents a CamK peptide fragment.

XX Sequence 22 AA;

Query Match 35.0%; Score 43; DB 21; Length 22;  
Best Local Similarity 66.7%; Pred. No. 4.7;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 11 RTEADCKAEN 22

Db 3 rqaavdclkkfn 14

RESULT 3

AAP80554

ID AAP80554 standard; protein; 19 AA.

XX AC AAP80554;

DT 28-OCT-1990 (first entry)

DE Partial sequence of bovine growth hormone.

XX Bovine growth hormone; peptide; partial sequence; antibody.

OS Bos taurus.

XX EP284406-A.

PD 28-SEP-1988.

PF 25-MAR-1988; 88EP-0302656.

PR 27-MAR-1987; 87GB-0007398.

XX (COOP-) COOPERS ANIMAL HEALTH LTD.

PI James S, Aston R, Bomford R;

XX WPI; 1988-272829/39.

PT Small peptide(s) contg. bovine growth hormone partial sequence - used for raising antibodies which will potentiate activity of hormone in vertebrate.

PS Claim 2; Page 12; 23pp; English.

CC The peptide has primary structural homology to a bovine growth hormone sequence. It can be used to raise antibodies which will potentiate the activity of the hormone in a vertebrate. The peptide can also be used to treat an (ab)normal vertebrate to boost growth, to bring abnormally low levels of growth up to the norm, to boost milk yield, or to boost or enhance other biological effects of growth hormone. Proportion of lean meat to fat may also be enhanced. The peptide may be linked to other antigens to produce a dual effect, eg all/part of somatostatin to create anti-somatostatin antibodies which would promote growth, or to all/part of a sex hormone molecule to provide simultaneous immunocastration.

XX Sequence 19 AA;

Query Match 30.1%; Score 37; DB 9; Length 19;

Best Local Similarity 58.3%; Pred. No. 34;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYSISRTEAADC 17

Db 8 rysiqntqvafc 19

RESULT 4

AAAP90941

ID AAP90941 standard; peptide; 19 AA.

XX AC AAP90941;

DT 02-MAR-1990 (first entry)

DE Residues 35-53 of bovine growth hormone.

XX Foot and mouth disease virus; bovine growth hormone; antibodies.

OS Bovidae.

PN WO8909228-A.

XX 05-OCT-1989.

XX

```

PF 23-MAR-1989; 89WO-GB00311.
XX
PR 25-MAR-1988; 88EP-0302656.
PR 08-SEP-1988; 88GB-0021076.
XX
XX (COOP-) COOPERS ANIMAL HEALTH LTD.
XX (WELL ) THE WELLCOME FOUNDATION LTD.
XX James S, Rowlands DJ, Francis MJ;
XX WPI; 1989-309504/42.
XX
XX New peptide contg. T-helper cell epitope of foot and mouth virus - and
XX opt. B-cell epitope, useful in vaccines and for potentiating hormone
XX activity.
XX
XX Claim 9; page 25; 31pp; English.
XX
XX The peptide is residues 35-53 of bovine growth hormone. It can be used
XX with peptides from foot and mouth disease virus in a vaccine to prevent
XX FMDV infection, or to potentiate hormone activity, eg to improve growth
XX rate or milk prodn.
XX
XX Sequence 19 AA;

Query Match 30.1%; Score 37; DB 10; Length 19;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYSISRTEAADC 17
DB 8 rysiqtqvafc 19
    |||| | : | |
    |||| | : | |

RESULT 5
AAR49699
ID AAR49699 standard; peptide; 19 AA.
XX
XX AAR49699;
XX
XX 30-JUL-1994 (first entry)
XX
XX Cattle, sheep GH peptide 35-53.
XX
XX Growth hormone; GH; somatotropin; growth hormone releasing factor;
XX GRF; somatotliberin; somatostatin; ruminant; anabolism;
XX feed-additive; immunization; vaccine.
XX
XX Bos taurus; Ovis aries.
XX
XX WO9404187-A.
XX
XX 03-MAR-1994.
XX
XX 04-AUG-1993; 93WO-FR00793.
XX
XX 20-AUG-1992; 92FR-0010159.
XX
XX (INMR ) RHONE MERIEUX SA.
XX
XX Dufour R, Roulet C;
XX
XX WPI; 1994-082840/10.
XX
XX Improving growth and carcass quality in pigs and ruminants - by
XX stimulating secretion of endogenous growth hormone, partic. by
XX vaccination with e.g. somatostatin
XX
XX Disclosure; Page 29; 48pp; French.
XX
XX GH secretion in ruminants is induced by anti-somatostatin
XX immunization and/or by potentiating immunomodulation of GRF using 1

CC or more peptide fragments of GH, especially those of sequence
CC AAR49699-719 and R455776-77, and/or peptide fragments of GRF.
XX
XX Sequence 19 AA;

Query Match 30.1%; Score 37; DB 15; Length 19;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYSISRTEAADC 17
DB 8 rysiqtqvafc 19
    |||| | : | |
    |||| | : | |

RESULT 6
AAP80556
ID AAP80556 standard; protein; 19 AA.
XX
XX AAP80556;
XX
XX 28-OCT-1990 (first entry)
XX
XX Partial sequence of bovine growth hormone.
XX
XX Bovine growth hormone; peptide; partial sequence; antibody.
XX
XX Bos taurus.
XX
XX EP284406-A.
XX
XX 28-SEP-1988.
XX
XX 25-MAR-1988; 88EP-0302656.
XX
XX 27-MAR-1987; 87GB-0007398.
XX
XX (COOP-) COOPERS ANIMAL HEALTH LTD.
XX
XX James S, Aston R, Bomford R;
XX WPI; 1988-272829/39.
XX
XX Small peptide(s) contg. bovine growth hormone partial sequence - used
XX for raising antibodies which will potentiate activity of hormone
XX in vertebrate.
XX
XX Claim 2; Page 12; 23pp; English.
XX
XX The peptide has primary structural homology to a bovine growth hormone
XX sequence. It can be used to raise antibodies which will potentiate the
XX activity of the hormone in a vertebrate. The peptide can also be used to
XX treat an (ab)normal vertebrate to boost growth, to bring abnormally
XX low levels of growth up to the norm, to boost milk yield, or to boost or
XX enhance other biological effects of growth hormone. Proportion of lean
XX meat to fat may also be enhanced. The peptide may be linked to other
XX antigens to produce a dual effect, eg all/part of somatostatin to create
XX anti-somatostatin antibodies which would promote growth, or to all/part
XX of a sex hormone molecule to provide simultaneous immunocastration.
XX
XX Sequence 19 AA;

Query Match 29.3%; Score 36; DB 9; Length 19;
Best Local Similarity 58.3%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYSISRTEAADC 17
DB 8 rysiqtqvafc 19
    |||| | : | |
    |||| | : | |

RESULT 7

```



AAP90943  
ID AAP90943 standard; peptide; 19 AA.

XX AC AAP90943;  
XX DT 02-MAR-1990 (first entry)  
XX DE Variant of residues 35-53 of bovine growth hormone.  
XX KW Foot and mouth disease virus; bovine growth hormone; antibodies.  
XX OS Bovidae.  
XX PN W08909228-A.  
XX PD 05-OCT-1989.  
XX PF 23-MAR-1989; 89WO-GB00311.  
XX PR 25-MAR-1988; 88EP-0302656.  
XX PR 08-SEP-1988; 88GB-0021076.  
XX PA (COOP-) COOPERS ANIMAL HEALTH LTD.  
XX PA (WELL) THE WELLCOME FOUNDATION LTD.  
XX PI James S, Rowlands DJ, Francis MJ;  
XX WPI; 1989-309504/42.  
XX DR New peptide contg. T-helper cell epitope of foot and mouth virus - and  
XX PT opt. B-cell epitope, useful in vaccines and for potentiating hormone  
XX PT activity.  
XX PS Claim 9; page 25; 31pp; English.  
XX CC The peptide is a variant which is antigenically equiv. to residues 35-53  
XX CC of bovine growth hormone. It can be used with peptides from foot and  
XX CC mouth disease virus in a vaccine to prevent FMDV infection, or to  
XX CC potentiate hormone activity, eg to improve growth rate or milk prodn.  
XX SQ Sequence 19 AA;

Query Match 29.3%; Score 36; DB 10; Length 19;  
Best Local Similarity 58.3%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYISIRTEAADC 17  
| | | | : | | |  
Db 8 rysiqaqaafc 19

RESULT 8  
AAR49700  
ID AAR49700 standard; peptide; 19 AA.  
XX AC AAR49700;  
XX DT 30-JUL-1994 (first entry)  
XX DE Pig GH peptide 35-53.  
XX KW Growth hormone; GH; somatotropin; growth hormone releasing factor;  
XX KW GRF; somatoliberin; somatostatin; ruminant; anabolism;  
XX KW feed-additive; immunization; vaccine.  
XX OS Sus scrofa.  
XX PN W09404187-A.  
XX PD 03-MAR-1994.  
XX PR 04-AUG-1993; 93WO-FR00793.

Query Match 29.3%; Score 36; DB 10; Length 19;  
Best Local Similarity 58.3%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYISIRTEAADC 17  
| | | | : | | |  
Db 8 rysiqaqaafc 19

RESULT 8  
AAR49700  
ID AAR49700 standard; peptide; 19 AA.  
XX AC AAR49700;  
XX DT 30-JUL-1994 (first entry)  
XX DE Pig GH peptide 35-53.  
XX KW Growth hormone; GH; somatotropin; growth hormone releasing factor;  
XX KW GRF; somatoliberin; somatostatin; ruminant; anabolism;  
XX KW feed-additive; immunization; vaccine.  
XX OS Sus scrofa.  
XX PN W09404187-A.  
XX PD 03-MAR-1994.  
XX PR 04-AUG-1993; 93WO-FR00793.

XX 20-AUG-1992; 92FR-0010159.  
XX (INMR) RHONE MERIEUX SA.  
XX PI Dufour R, Roulet C;  
XX WPI; 1994-082840/10.  
XX Improving growth and carcass quality in pigs and ruminants - by  
XX stimulating secretion of endogenous growth hormone, partic. by  
XX vaccination with e.g. somatostatin  
XX Disclosure; Page 29; 48pp; French.  
XX GH secretion in ruminants is induced by anti-somatostatin  
XX immunization and/or by potentiating immunomodulation of GRF using 1  
XX or more peptide fragments of GH, especially those of sequence  
XX AAR49699-719 and R455776-77, and/or peptide fragments of GRF.  
XX SQ Sequence 19 AA;

Query Match 29.3%; Score 36; DB 15; Length 19;  
Best Local Similarity 58.3%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYISIRTEAADC 17  
| | | | : | | |  
Db 8 rysiqaqaafc 19

RESULT 9  
AAW35265  
ID AAW35265 standard; peptide; 21 AA.  
XX AC AAW35265;  
XX DT 17-FEB-1998 (first entry)  
XX DE Epitope comprising porcine somatotropin 1-190 residues 33-53.  
XX KW Epitope; pig; porcine; somatotropin; somatotrophin;  
XX KW recombinant production; composite peptide; immunisation;  
XX KW antibody.  
XX OS Sus scrofa.  
XX PN US5686268-A.  
XX PD 11-NOV-1997.  
XX PF 27-JAN-1995; 95US-0388267.  
XX PR 19-JUN-1992; 92US-0901704.  
XX PR 27-JAN-1995; 95US-0388267.  
XX PA (PFIZ) PFIZER INC.  
XX PI Alila HW, Clark MT, Jones EV, Miller TJ, O'Brien SP;  
XX PI Sathe GM;  
XX WPI; 1997-558137/51.  
XX DNA for composite peptide comprising at least 2 somatotropin  
XX epitope(s) - useful as immunogen to elicit somatotropin stabilising  
XX antibodies  
XX Disclosure; Columns 29-30; 32pp; English.  
XX The present sequence is an epitope comprising porcine somatotropin  
XX (pST) 1-190 residues 33-53. An expression vector containing a DNA  
XX molecule encoding a composite peptide, which comprises at least 2

Query Match 29.3%; Score 36; DB 15; Length 19;  
Best Local Similarity 58.3%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYISIRTEAADC 17  
| | | | : | | |  
Db 8 rysiqaqaafc 19

RESULT 9  
AAW35265  
ID AAW35265 standard; peptide; 21 AA.  
XX AC AAW35265;  
XX DT 17-FEB-1998 (first entry)  
XX DE Epitope comprising porcine somatotropin 1-190 residues 33-53.  
XX KW Epitope; pig; porcine; somatotropin; somatotrophin;  
XX KW recombinant production; composite peptide; immunisation;  
XX KW antibody.  
XX OS Sus scrofa.  
XX PN US5686268-A.  
XX PD 11-NOV-1997.  
XX PF 27-JAN-1995; 95US-0388267.  
XX PR 19-JUN-1992; 92US-0901704.  
XX PR 27-JAN-1995; 95US-0388267.  
XX PA (PFIZ) PFIZER INC.  
XX PI Alila HW, Clark MT, Jones EV, Miller TJ, O'Brien SP;  
XX PI Sathe GM;  
XX WPI; 1997-558137/51.  
XX DNA for composite peptide comprising at least 2 somatotropin  
XX epitope(s) - useful as immunogen to elicit somatotropin stabilising  
XX antibodies  
XX Disclosure; Columns 29-30; 32pp; English.  
XX The present sequence is an epitope comprising porcine somatotropin  
XX (pST) 1-190 residues 33-53. An expression vector containing a DNA  
XX molecule encoding a composite peptide, which comprises at least 2

CC non-contiguous ST epitopic amino acid sequences, e.g. the present  
 CC sequence, and is free of receptor binding domain sequences, can be  
 CC used to transform a host cell. The transformant cell can be used for  
 CC the recombinant production of the composite peptide, which can be  
 CC used as an immunogen to stimulate the production of anti-pST  
 CC antibodies (Ab) in pigs. Such Ab may increase the half life of  
 CC co-administered pST in circulation.

XX Sequence 21 AA;

Query Match 29.3%; Score 36; DB 18; Length 21;  
 Best Local Similarity 58.3%; Pred. No. 54;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYSISRTEAAC 17  
 | | | | : | | |  
 Db 9 rysiqaqaafc 20

#### RESULT 10

AAY05067  
 ID AAY05067 standard; peptide; 19 AA.

XX  
 AC AAY05067;

DT 16-JUN-1999 (first entry)

DE Tumour antigen antibody heavy chain CDR3 clone TA50.

KW Tumour antigen; antibody; CDR; complementarity determining region;  
 KW binding molecule identification; tumour-specific binding polypeptide;  
 KW cancer therapy; heavy chain.

XX Homo sapiens.

OS WO9906834-A2.

PN 11-FEB-1999.

PD 04-AUG-1998; 98WO-US16280.

PF 04-AUG-1997; 97US-0905825.

PR 04-AUG-1997; 97US-1112222.

XX (IXSY-) IXSYS INC.

PI Huse, WD, Watkins JD, Wu H;

XX WPI; 1999-153951/13.

DR N-PSDB; AAX28244.

XX Identifying binding molecules for ligands, particularly tumour  
 PT antigens - by selectively immobilising a population of binding  
 PT molecules to a solid support and screening for binding to two or  
 PT more ligands

PS Claim 15; Page 61; 80pp; English.

XX This sequence represents a heavy chain complementarity determining  
 CC region (CDR) from a tumour antigen specific antibody.  
 CC The invention relates to a method for identifying a binding molecule  
 CC having selective affinity for a ligand comprising: (a) selectively  
 CC immobilising a diverse population of binding molecules to a solid  
 CC support; (b) simultaneously contacting the diverse population immobilised  
 CC on the solid support with 2 or more ligands; and (c) determining at least  
 CC one binding molecule which selectively binds to one or more of the  
 CC ligands. The method allows for the rapid and efficient methods for the  
 CC identification of binding molecules which exhibit selective affinity for  
 CC one or more ligands of interest. They are used particularly for  
 CC identifying tumour-specific binding polypeptides which can be used as  
 CC targeting agents for cancer therapy that minimises impact on non-tumour  
 CC tissues.

XX Sequence 19 AA;

Query Match 28.5%; Score 35; DB 20; Length 19;  
 Best Local Similarity 46.2%; Pred. No. 70;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 KNGRYSISRTEAA 15  
 | | | | : | | |  
 Db 3 ktsrytlrrseas 15

#### RESULT 11

AAR96887  
 ID AAR96887 standard; peptide; 15 AA.

XX  
 AC AAR96887;

DT 29-NOV-1996 (first entry)

DE Human cartilage link protein derivative H1.

XX IgA protease precursor; Ipp; bacterial polypeptide; autoimmunity;  
 KW viral infection; rheumatoid arthritis; AIDS; meningococcal bacteria;  
 KW immunoglobulin; human connective tissue; cartilage; link protein;  
 KW proteoglycan tandem repeat; PTR; Neisseria gonorrhoeae; MHC;  
 KW major histocompatibility complex; class II; self epitope; autoantigen.

OS Homo sapiens.

XX WO9609395-A2.

PN 28-MAR-1996.

PD 21-SEP-1995; 95WO-EP03726.

PF 21-SEP-1994; 94DE-4433708.

PR (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PI Beck SC, Jose J, Lorenzen DR, Meyer TFF, Oetzelberger KB;  
 PI Pohlner J, Woelk U;

XX WPI; 1996-188456/19.

XX Medicaments for treating auto-immune or viral diseases - contg.  
 PT substances interfering with bacterial poly:protein function

PS Claim 32; Page 86; 117pp; German.

XX The present sequence is from human cartilage link protein which  
 CC assembles into large aggregates with hyaluronan and proteoglycan  
 CC core protein. The region is homologous to the proteoglycan tandem  
 CC repeat sequences found in many connective tissue proteins; it is  
 CC also homologous to the gamma peptide of the precursor of IgA-protease  
 CC polyprotein (Ipp) of pathogenic Neisseria gonorrhoeae strains MS11  
 CC and K16. The Neisseria Ipp has marked homology to certain human  
 CC proteins and has been implicated in rheumatoid arthritis and other  
 CC auto-immune diseases. The polyprotein also activates proviruses,  
 CC including HIV. Substances which interfere with the function of Ipp  
 CC from Neisseria will be useful for treating associated autoimmune  
 CC diseases and viral infections. Four peptides (designated H1-H4,  
 CC see AAR96887-R96890) were derived from the link protein and were  
 CC assayed for binding to MHC alleles DR1Dw1, DR4Dw4 and DR3. Peptide  
 CC H1 did not bind to any of the alleles.

XX Sequence 15 AA;

Query Match 27.6%; Score 34; DB 17; Length 15;  
 Best Local Similarity 42.9%; Pred. No. 77;  
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 GRYSISRTEADCC 18  
 |||::: || |  
 Db 1 grynlnfheaqqac 14

## RESULT 12

AAG62706  
 ID AAG62706 standard; peptide; 17 AA.

XX AAG62706;

DT 17-SEP-2001 (first entry)

XX Hanatoxin-like sequence from type 3 semaphorin hSema3C.

XX Hanatoxin; tarantula; toxin; voltage-gated potassium channel;

KW voltage-gated calcium channel; hanatoxin-like sequence; HTLS;

KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.

XX Homo sapiens.

XX WO200138491-A2.

XX 31-MAY-2001.

PF 07-NOV-2000; 2000WO-US41943.

XX 08-NOV-1999; 99US-0164056.

XX (GEO) GEN HOSPITAL CORP.

PI Behar O, Woolf CJ;

XX WPI; 2001-451494/48.

XX Polypeptide sequences that encompass the hanatoxin-like sequences of  
 semaphorins, useful as a drugs to treat any condition or disease that  
 is characterized by abnormal calcium channel function -

XX Claim 3; Page 10; 29pp; English.

XX AAG62702-22 represent hanatoxin-like sequences (HTLS) found in  
 the semaphorin domain of mammalian secreted semaphorins. Hanatoxin  
 is a tarantula toxin that selectively blocks some voltage-gated  
 potassium and calcium channels. The HTLS is responsible for the dorsal  
 root ganglion repulsion and growth cone collapse activities associated  
 with semaphorins. Polypeptides containing HTLS can be used to modulate  
 the activity of calcium channels. The peptides can also be used as an  
 antigen to generate antibodies that can then be used to modulate the  
 activity of calcium channels by inactivating naturally occurring channel  
 ligands. The peptides or antibodies can be used as drugs to treat any  
 condition or disease that is characterized by abnormal calcium channel  
 function.

XX Sequence 17 AA;

Query Match 27.6%; Score 34; DB 22; Length 17;  
 Best Local Similarity 66.7%; Pred. No. 88;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 12 TEAADCCCA 20

Db 1 tacadccla 9

## RESULT 13

AAW35266  
 ID AAW35266 standard; peptide; 19 AA.

XX AAW35266;

DT 17-FEB-1998 (first entry)

XX Epitope comprising porcine somatotropin 1-191 residues 35-53.

XX Epitope; pig; porcine; somatotropin; somatotrophin;

KW recombinant production; composite peptide; immunisation;  
 antibody.

XX Sus scrofa.

XX US5686268-A.

XX 11-NOV-1997.

XX 27-JAN-1995; 95US-0388267.

XX 19-JUN-1992; 92US-0901704.

PR 27-JAN-1995; 95US-0388267.

XX (PFIZ) PFIZER INC.

XX Allila HW, Clark MT, Jones EV, Miller TJ, O'Brien SP;  
 PI Sathe GM;

XX WPI; 1997-558137/51.

XX DNA for composite peptide comprising at least 2 somatotropin  
 epitope(s) - useful as immunogen to elicit somatotropin stabilising  
 antibodies

XX Disclosure; Columns 31-32; 32pp; English.

XX The present sequence is an epitope comprising porcine somatotropin  
 (pST) 1-191 residues 35-53. An expression vector containing a DNA  
 molecule encoding a composite peptide, which comprises at least 2  
 non-contiguous ST epitopic amino acid sequences, e.g. the present  
 sequence, and is free of receptor binding domain sequences, can be  
 used to transform a host cell. The transformed cell can be used for  
 the recombinant production of the composite peptide, which can be  
 used as an immunogen to stimulate the production of anti-pST  
 antibodies (Ab) in pigs. Such Ab may increase the half life of  
 co-administered pST in circulation.

XX Sequence 19 AA;

Query Match 27.6%; Score 34; DB 18; Length 19;  
 Best Local Similarity 58.3%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 RYSISRTEADCC 17

Db 8 rysiqnanaafc 19

## RESULT 14

AAAY99910  
 ID AAAY99910 standard; Peptide; 13 AA.

XX AAAY99910;

XX 16-NOV-2000 (first entry)

XX Peptide encoded by Snut 2425ES DNA used in HIV DNA vaccine.

XX HIV; human immunodeficiency virus; vaccine; AIDS; snut;  
 KW silent nucleotide substitution.

XX Human immunodeficiency virus type 1.

XX WO200029561-A2.

XX 25-MAY-2000.

XX 27-MAR-2000; 2000WO-DK00144.  
 XX  
 PF  
 XX  
 PR 29-MAR-1999; 99DK-0000427.  
 PR 09-APR-1999; 99US-0128558.  
 XX  
 XX (STAT-) STATENS SERUM INST.  
 XX PA  
 XX Fomsgaard A;  
 XX PI  
 XX WPI; 2000-387778/33.  
 DR N-PSDB; AAA49079.

XX Producing nucleotide sequence construct with optimized codons for human  
 XX immunodeficiency virus (HIV) genetic vaccine involves obtaining a first  
 PT nucleotide sequence from a HIV patient, redesigning and assembling it  
 PT with snuts -  
 XX

PS Example 3; 107; 150pp; English.

XX The present invention relates to a nucleotide construct with optimised  
 CC codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The  
 CC construct uses codons from highly expressed mammalian proteins to code  
 CC for each derivative of an early, primary HIV envelope gene. The first  
 CC stage in the production of the construct was the cloning of an HIV  
 CC envelope gene. A nucleotide sequence encoding this gene was then created  
 CC using codons from highly expressed mammalian genes. The present sequence  
 CC is the peptide encoded by one of the snuts (AAA49060-A49079) that were  
 CC created by redesigning the nucleotide construct so that restriction  
 CC enzyme sites surrounded functional regions of the sequence. The snuts  
 CC were then assembled into pieces (AAA49080-A49092). Each derivative of  
 CC the envelope gene (AAA49093-A49097) was then built using the pieces.  
 CC The HIV DNA vaccine may be used as a prophylactic vaccine and as a  
 CC therapeutic vaccine in HIV infected patients.  
 XX

SQ Sequence 13 AA;

Query Match 26.8%; Score 33; DB 21; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 94;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 SRTEADCK 19  
 :| :| :| :|  
 Db 3 arassapck 12

RESULT 15

AAAY97089  
 ID AAY97089 standard; peptide; 20 AA.

XX AAY97089;

DT 22-DEC-2000 (first entry)

XX Tau conotoxin sequence.

XX Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;  
 KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;  
 KW migraine; treatment; therapy.

XX Conus marmoreus.

XX Key Location/Qualifiers

FT Misc-difference 14 /note= "Glu or gamma-carboxy-Glu"

FT Misc-difference 16

FT /note= "Lys, N-methyl-Lys, N,N-dimethyl-Lys or  
 N,N,N-trimethyl-Lys"

FT Misc-difference 17

FT /note= "Glu or gamma-carboxy-Glu"

XX WO2000046371-A1.

XX 10-AUG-2000.  
 XX PD  
 XX PF  
 XX 04-FEB-2000; 2000WO-US03021.  
 XX  
 XX 04-FEB-1999; 99US-0118642.  
 XX  
 XX (UTAH ) UNIV UTAH RES FOUND.  
 XX (COGN-) COGNETIX INC.  
 XX Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;  
 PI Jones RM;  
 PI WPI; 2000-543489/49.

XX Novel analgesic short peptides from predatory cone snails for treating  
 PT acute, chronic and neuropathic pain and migraines  
 PT  
 XX Claim 2; Page 28; 58pp; English.

XX The venom of predatory cone snails comprises relatively small  
 CC peptides which are targeted to various neuromuscular receptors  
 CC and may be equivalent in their pharmacological diversity to the  
 CC alkaloids of plants or secondary metabolites of microorganisms.  
 CC Several peptides have been characterised from Conus venoms. These  
 CC include the alpha and mu conotoxins which target nicotinic  
 CC acetylcholine receptors and muscic sodium channels respectively.  
 CC Chronic or intractable pain as well as neuropathic pain is  
 CC currently treated with a range of analgesic compounds. It is  
 CC thought that the tau-conotoxins described could have applications  
 CC as analgesic drugs and could be used for treating such acute,  
 CC chronic and neuropathic pain as well as migraines.  
 XX

SQ Sequence 20 AA;

Query Match 26.8%; Score 33; DB 21; Length 20;  
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 AADCCAKF 21  
 :| :| :| :| :|  
 Db 6 agdcccrrf 13

Search completed: February 21, 2002, 16:26:41  
 Job time: 32 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:26:14 ; Search time 12.47 Seconds  
(without alignments)  
39.701 Million cell updates/sec

Title: US-08-753-851-1  
Perfect score: 123  
Sequence: 1 CEKNGRYSISRTEADCKAFN 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 117372

Minimum DB seq length: 0  
Maximum DB seq length: 22

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents,AA:\*  
1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	100.0	22	2 US-08-143-311B-1	Sequence 1, Appli
2	37	30.1	19	2 US-07-921-447-3	Sequence 3, Appli
3	36	29.3	19	2 US-07-921-447-5	Sequence 5, Appli
4	36	29.3	21	1 US-08-388-267C-5	Sequence 6, Appli
5	34	27.6	19	1 US-08-388-267C-6	Sequence 27, Appli
6	32	26.0	16	2 US-08-484-905-27	Sequence 27, Appli
7	32	26.0	16	3 US-08-481-985B-27	Sequence 27, Appli
8	32	26.0	16	4 US-08-370-476-27	Sequence 26, Appli
9	32	26.0	19	3 US-09-023-082A-26	Sequence 132, App
10	31	25.2	7	4 US-09-139-802-132	Sequence 48, Appli
11	31	25.2	20	4 US-08-290-736C-48	Sequence 5, Appli
12	31	25.2	21	1 US-07-696-051B-5	Sequence 33, Appli
13	30	24.4	8	4 US-08-481-968A-33	Sequence 31, Appli
14	30	24.4	15	1 US-08-488-212A-31	Sequence 31, Appli
15	30	24.4	15	2 US-08-320-306-31	Sequence 31, Appli
16	30	24.4	15	2 US-08-488-209B-31	Sequence 31, Appli
17	30	24.4	15	2 US-08-408-011-31	Sequence 14, Appli
18	30	24.4	15	3 US-08-466-860-14	Sequence 14, Appli
19	30	24.4	15	3 US-08-472-040A-14	Sequence 14, Appli
20	30	24.4	15	4 US-08-276-776-14	Sequence 14, Appli
21	30	24.4	15	4 US-08-471-209-14	Sequence 14, Appli
22	30	24.4	18	6 5210180-5	Patent No. 5210180
23	30	24.4	22	1 US-08-221-580-3	Sequence 3, Appli
24	30	24.4	22	5 PCT-US95-04018-65	Sequence 65, Appli
25	29	23.6	6	4 US-09-329-350-5	Sequence 5, Appli
26	29	23.6	7	4 US-09-329-350-8	Sequence 8, Appli
27	29	23.6	9	4 US-09-189-060B-57	Sequence 57, Appli

Sequence 236, App  
Sequence 102, App  
Sequence 103, App  
Sequence 104, App  
Sequence 108, App  
Sequence 79, Appli  
Sequence 3, Appli  
Sequence 45, Appli  
Sequence 27, Appli  
Sequence 28, Appli  
Sequence 26, Appli  
Sequence 23, Appli  
Sequence 8, Appli  
Sequence 98, Appli  
Sequence 17, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 11, Appli

## ALIGNMENTS

RESULT 1  
US-08-143-311B-1  
; Sequence 1, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHUYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-1



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: PFIZER INC.  
;; STREET: 235 EAST 42ND STREET  
;; CITY: NEW YORK  
;; STATE: NY  
;; COUNTRY: U.S.A.  
;; ZIP: 10017-5755  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/388,267C  
;; FILING DATE: 27-JAN-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: KOLLER, ALAN L.  
;; REGISTRATION NUMBER: 37,371  
;; REFERENCE/DOCKET NUMBER: PC9042A  
;; TELEPHONE: 212-573-2118  
;; TELEFAX: 212-808-8893  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; ORIGINAL SOURCE:  
;; ORGANISM: Sus scrofa  
;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 1..21  
;; OTHER INFORMATION: /note= "REPRESENTS RESIDUES 33-53  
;; OF PST1-190"  
US-08-388-267C-5

Query Match 29.3%; Score 36; DB 1; Length 21;  
Best Local Similarity 58.3%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RYSISRTEAACD 17  
||||| :|||  
DB 9 RYSIQNAQAFC 20

RESULT 5  
US-08-388-267C-6  
;; Sequence 6, Application US/08388267C  
;; Patent No. 5686268  
;; GENERAL INFORMATION:  
;; APPLICANT: JONES, ELAINE V.  
;; APPLICANT: SATHE, GANESH M.  
;; APPLICANT: O'BRIEN, SHAWN  
;; APPLICANT: CLARK, MICHAEL T.  
;; APPLICANT: ALILA, HECTOR W.  
;; APPLICANT: MILLER, TIMOTHY J.  
;; TITLE OF INVENTION: FUSION PROTEINS  
;; NUMBER OF SEQUENCES: 24  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: PFIZER INC.  
;; STREET: 235 EAST 42ND STREET  
;; CITY: NEW YORK  
;; STATE: NY  
;; COUNTRY: U.S.A.  
;; ZIP: 10017-5755  
;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/388,267C  
;; FILING DATE: 27-JAN-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: KOLLER, ALAN L.  
;; REGISTRATION NUMBER: 37,371  
;; REFERENCE/DOCKET NUMBER: PC9042A  
;; TELEPHONE: 212-573-2118  
;; TELEFAX: 212-808-8893  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 19 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; ORIGINAL SOURCE:  
;; ORGANISM: Sus scrofa  
;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 1..19  
;; OTHER INFORMATION: /note= "THIS PEPTIDE REPRESENTS  
;; OTHER INFORMATION: RESIDUES 35-53 OF PST1-191"  
US-08-388-267C-6

Query Match 27.6%; Score 34; DB 1; Length 19;  
Best Local Similarity 58.3%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 RYSISRTEAACD 17  
||||| :|||  
DB 8 RYSIQNAQAFC 19

RESULT 6  
US-08-484-905-27  
;; Sequence 27, Application US/08484905  
;; Patent No. 5976551  
;; GENERAL INFORMATION:  
;; APPLICANT: Mottez, Estelle  
;; APPLICANT: Abastado, Jean-Pierre  
;; APPLICANT: Kourilsky, Philippe

;; TITLE OF INVENTION: An Altered Major Histocompatibility  
;; Determinant  
;; TITLE OF INVENTION: Determinant  
;; NUMBER OF SEQUENCES: 127  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
;; ADDRESSEE: Dunner  
;; STREET: 1300 I Street, N.W., Suite 700  
;; CITY: Washington  
;; STATE: D.C.  
;; ZIP: 20005-3315  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy Disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS-/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/484,905  
;; FILING DATE: 07-JUNE-1995  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:





US-08-370-476-27

Query Match 26.0%; Score 32; DB 4; Length 16;  
Best Local Similarity 41.7%; Pred. No. 78;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 KNGRYSISRTEA 14  
:||: :||:|  
Db 4 ENKETLQRTDA 15

## RESULT 9

US-09-023-082A-26  
; Sequence 26, Application US/09023082A  
; Patent No. 6077692  
; GENERAL INFORMATION:  
; APPLICANT: RUBEN, STEVEN M.  
; APPLICANT: JIMENEZ, PABLO  
; APPLICANT: DUAN, D. ROXANNE  
; APPLICANT: RAMPY, MARK A.  
; APPLICANT: MENDRICK, DONNA  
; APPLICANT: ZHANG, JUN  
; APPLICANT: NI, JIAN  
; APPLICANT: MOORE, PAUL A.  
; APPLICANT: COLEMAN, TIMOTHY A.  
; APPLICANT: GRUBER, JOACHIM R.  
; APPLICANT: DILLON, PATRICK J.  
; APPLICANT: GENTZ, REINER L.  
; TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR-2  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
; STREET: 1100 NEW YORK AVE, NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,082A  
; FILING DATE: 13-FEB-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01790  
; FILING DATE: 14-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/461,195  
; FILING DATE: 05-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/023,852  
; FILING DATE: 13-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/039,045  
; FILING DATE: 28-FEB-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/862,432  
; FILING DATE: 23-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/910,875  
; FILING DATE: 13-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,561  
; FILING DATE: 13-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.0360008/EKS  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; US-09-023-082A-26

Query Match 26.0%; Score 32; DB 3; Length 19;  
Best Local Similarity 43.8%; Pred. No. 94;  
Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 EKNGRYSISRTEAACD 17  
||||: :||:|  
Db 3 EKNGK--VSGTKKENC 16

## RESULT 10

US-09-139-802-132  
; Sequence 132, Application US/09139802  
; Patent No. 6180084  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing  
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: P-LJ 3203  
; CURRENT APPLICATION NUMBER: US/09/139,802  
; CURRENT FILING DATE: 1998-08-25  
; EARLIER APPLICATION NUMBER: 08/926,914  
; EARLIER FILING DATE: 1997-09-10  
; EARLIER APPLICATION NUMBER: 08/710,067  
; EARLIER FILING DATE: 1996-09-10  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 132  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
; US-09-139-802-132

Query Match 25.2%; Score 31; DB 4; Length 7;  
Best Local Similarity 66.7%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 CCKAFN 22  
||:|  
Db 2 CCRQFN 7

## RESULT 11

US-08-290-736C-48  
; Sequence 48, Application US/08290736C  
; Patent No. 6294174  
; GENERAL INFORMATION:  
; APPLICANT: KRSMANOVIC, VELIBOR  
; COSIC, IRENA  
; BLOQUARD, JEAN-MICHEL  
; HEARN, MILTON TW  
; TITLE OF INVENTION: PEPTIDES IMMUNOLOGICALLY RELATED TO  
; PROTEINS OF A VIRAL AGENT AND THEIR BIOLOGICAL APPLICATI  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.



Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 14 AADCKA 20  
||:||||  
Db 1 AANCAA 7

## RESULT 14

US-08-488-212A-31  
; Sequence 31, Application US/08488212A  
; Patent No. 5665355  
; GENERAL INFORMATION:  
; APPLICANT: Primi, Daniele  
; TITLE OF INVENTION: Diagnosis and Treatment of  
; TITLE OF INVENTION: AIDS Onset  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Thomas E. Popovich, Thomas  
; ADDRESSEE: Popovich & Associates  
; STREET: 80 South 8th Street  
; CITY: Minneapolis  
; STATE: Minnesota  
; COUNTRY: USA  
; ZIP: 55402-2111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB  
; MEDIUM TYPE: Storage  
; COMPUTER: IBM Compatible Compaq Prolinea  
; COMPUTER: 4/66  
; OPERATING SYSTEM: MS-DOS Version 5  
; SOFTWARE: Microsoft Word for Windows  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,212A  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,485  
; FILING DATE: No. 5665355ember 9, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Thomas E. Popovich  
; REGISTRATION NUMBER: 30099  
; REFERENCE/DOCKET NUMBER: 3678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 334-8991  
; TELEFAX: (612) 334-8994  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: peptide  
; IMMEDIATE SOURCE: Chemical Synthesis  
US-08-488-212A-31

Query Match 24.4%; Score 30; DB 1; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 YSISRTEAAD 16  
||:||||  
Db 3 YSVSRSKTED 12

## RESULT 15

US-08-320-306-31  
; Sequence 31, Application US/08320306  
; Patent No. 5891623  
; GENERAL INFORMATION:  
; APPLICANT: Primi, Daniele  
; TITLE OF INVENTION: Diagnosis and Treatment of  
; TITLE OF INVENTION: AIDS Onset

NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thomas E. Popovich, Thomas  
ADDRESSEE: Popovich & Associates  
STREET: 80 South 8th Street  
CITY: Minneapolis  
STATE: Minnesota  
COUNTRY: USA  
ZIP: 55402-2111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible Compaq Prolinea  
COMPUTER: 4/66  
OPERATING SYSTEM: MS-DOS Version 5  
SOFTWARE: Microsoft Word for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/320,306  
FILING DATE: 06-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,485  
FILING DATE: No. 5891623ember 9, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. Popovich  
REGISTRATION NUMBER: 30099  
REFERENCE/DOCKET NUMBER: 3678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (612) 334-8991  
TELEFAX: (612) 334-8994  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: peptide  
IMMEDIATE SOURCE: Chemical Synthesis  
US-08-320-306-31

Query Match 24.4%; Score 30; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 YSISRTEAAD 16  
||:||||  
Db 3 YSVSRSKTED 12

Search completed: February 21, 2002, 16:27:00  
Job time: 46 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:29:24 ; Search time 12.81 Seconds  
(without alignments)  
136.769 Million cell updates/sec

Title: US-08-753-851-2  
Perfect score: 135  
Sequence: 1 CNTSQDYTCFNASAPPEEDCTS 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4466

Minimum DB seq length: 0  
Maximum DB seq length: 23

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR68: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	20.7	18	2 S39845	2-aminobenzoate-Co
2	27	20.0	19	2 C48363	2-hydroxyglutaryl-
3	27	20.0	21	2 S47207	T-cell receptor J-
4	26	19.3	12	2 PT0228	Ig heavy chain CDR
5	26	19.3	13	2 A28953	alpha-conotoxin SI
6	26	19.3	16	2 A39109	hypothetical prote
7	26	19.3	16	2 PH1317	Ig heavy chain DJ
8	25.5	18.9	20	2 S08605	hypothetical prote
9	25	18.5	12	2 PH1458	T-cell receptor be
10	25	18.5	15	2 B43474	thrombospondin 2 -
11	25	18.5	16	2 A60839	neurokinin A homol
12	25	18.5	16	2 PH0763	T-cell receptor be
13	25	18.5	16	2 PH0759	T-cell receptor be
14	25	18.5	17	2 A42920	fatty acid ethyl e
15	25	18.5	18	2 B24867	scyllorhizin II -
16	25	18.5	20	2 S77989	cytochrome-c oxida
17	25	18.5	20	2 S77992	cytochrome-c oxida
18	25	18.5	22	2 C39800	calcium-activated
19	24.5	18.1	23	2 B42382	replication initia
20	24	17.8	10	2 PC2171	triacylglycerol li
21	24	17.8	12	2 S26544	T-cell receptor be
22	24	17.8	13	1 NTKNAS	alpha-conotoxin SI
23	24	17.8	18	2 A56798	dermatan sulfate p
24	24	17.8	19	2 A44379	alpha-conotoxin SI
25	24	17.8	19	2 A61377	endometrial secret
26	24	17.8	20	2 JT0410	bombyxin-IV chain
27	24	17.8	20	2 B60977	tubulin beta chain
28	24	17.8	20	2 D84716	hypothetical prote
29	24	17.8	22	2 PH1359	Ig heavy chain DJ

30 23.5 17.4 16 2 A59042 alpha-conotoxin Ep  
31 23.5 17.4 21 2 I65270 collagen alpha 1(I  
32 23 17.0 8 2 S29272 tocopherol-binding  
33 23 17.0 12 2 C39109 hypothetical 1.2K  
34 23 17.0 13 2 A46463 preabsorbing antiq  
35 23 17.0 14 2 I54284 C1-inhibitor - hum  
36 23 17.0 15 2 PT0037 light harvesting c  
37 23 17.0 16 2 C39509 mannosyl-specific l  
38 23 17.0 17 2 A49237 45/47K antigen - M  
39 23 17.0 18 2 S24780 protein-tyrosine k  
40 23 17.0 19 2 PT0286 Ig heavy chain CDR  
41 23 17.0 19 2 S68393 H+-transporting AT  
42 23 17.0 19 2 S02808 nucleolin - bovine  
43 23 17.0 20 2 E60894 crystallin - Pacif  
44 23 17.0 20 2 C49404 T-cell receptor be  
45 23 17.0 23 2 A59048 convulsant peptide

## ALIGNMENTS

## RESULT 1

S39845  
2-aminobenzoate-CoA ligase, anaerobic - Pseudomonas sp. (strain KB740) (fragment)  
C:Species: Pseudomonas sp.  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
C:Accession: S39845  
R:Altenschmidt, U.; Fuchs, G.  
Eur. J. Biochem. 205, 721-727, 1992  
A:Title: Novel aerobic 2-aminobenzoate metabolism. Purification and characterization from a denitrifying Pseudomonas sp.  
A:Reference number: S22402; MUID:92241310  
A:Accession: S39845  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10, 'X', 12-14, 'X', 16-18 <ALT>

Query Match 20.7%; Score 28; DB 2; Length 18;  
Best Local Similarity 35.3%; Pred. No. 8.8e+02;  
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 3 TSQYDYTCFNASAPPEE 19

Db 1 TSHVDTFARRXLPPXEQ 17

## RESULT 2

C48363  
2-hydroxyglutaryl-CoA dehydratase gamma chain - Fusobacterium nucleatum (fragment)  
C:Species: Fusobacterium nucleatum  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995  
C:Accession: C48363  
R:Klees, A.G.; Linder, D.; Buckel, W.  
Arch. Microbiol. 158, 294-301, 1992  
A:Title: 2-Hydroxyglutaryl-CoA dehydratase from Fusobacterium nucleatum (subsp. nucle  
A:Reference number: A48363; MUID:93038007  
A:Accession: C48363  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-19 <KLE>  
A:Experimental source: subsp. nucleatum  
A:Note: sequence extracted from NCBI backbone (NCBIP:118485)

Query Match 20.0%; Score 27; DB 2; Length 19;  
Best Local Similarity 40.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NTSQDYTYCF 11

Db 2 NIXEYDDDFIF 11

## RESULT 3

S47207  
T-cell receptor J-alpha wNIII.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 05-Nov-1999  
C:Accession: S47207  
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.  
submitted to the EMBL Data Library, February 1993.  
A:Reference number: S40133  
A:Accession: S47207  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-21 <PLA>  
A:Cross-references: EMBL:X71032; NID:g506610; PIDN:CAA50349.1; PID:g510318  
C:Keywords: T-cell receptor

Query Match 20.0%; Score 27; DB 2; Length 21;  
Best Local Similarity 62.5%; Pred. No. 1.4e+03;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 SAPPEDC 21  
| | | | |  
DB 10 SLVPEDC 17

## RESULT 4

PT0228  
Ig heavy chain CDR3 region (clone 1-112) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0228  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0228  
A:Molecule type: DNA  
A:Residues: 1-12 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 19.3%; Score 26; DB 2; Length 12;  
Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
Matches 7; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 7 DTYCFNASAPPEED 20  
| | : : | | | |  
DB 1 DRYVROSS--PEED 12

## RESULT 5

A28953  
alpha-conotoxin SI - cone shell (Conus striatus)  
C:Species: Conus striatus (striated cone)  
C:Date: 30-Jun-1989 #sequence\_revision 25-Apr-1997 #text\_change 23-May-1997  
C:Accession: A28953  
R:Zafaralla, G.C.; Ramilo, C.; Gray, W.R.; Karlstrom, R.; Olivera, B.M.; Cruz, L.J.  
Biochemistry 27, 7102-7105, 1988  
A:Title: Phylogenetic specificity of cholinergic ligands: alpha-conotoxin SI.  
A:Reference number: A28953; MUID:89062448  
A:Accession: A28953  
A:Molecule type: protein  
A:Residues: 1-13 <ZAF>  
A:Note: This sequence was confirmed by chemical synthesis  
A:Comment: This paralytic toxin from a fish-hunting cone snail inhibits the acetylcholin  
C:Superfamily: alpha-conotoxin  
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurot  
F:2-7,3-13/Disulfide bonds: #status experimental  
F:13/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 19.3%; Score 26; DB 2; Length 13;  
Best Local Similarity 33.3%; Pred. No. 1.2e+03;  
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 10 CFNASAPPEEDC 21  
| | : : | |  
DB 2 CCNPACGPKYSC 13

## RESULT 6

A39109  
hypothetical protein 1 - hepatitis C virus  
C:Species: hepatitis C virus  
C:Date: 18-Oct-1991 #sequence\_revision 18-Oct-1991 #text\_change 18-Jun-1993  
C:Accession: A39109  
R:Han, J.H.; Shyamala, V.; Richman, K.H.; Brauer, M.J.; Irvine, B.; Urdea, M.S.; Teko  
Proc. Natl. Acad. Sci. U.S.A. 88, 1711-1715, 1991  
A:Title: Characterization of the terminal regions of hepatitis C viral RNA: identific  
A:Reference number: A39109; MUID:91156678  
A:Accession: A39109  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-16 <HAN>  
A:Cross-references: GB:M58406

Query Match 19.3%; Score 26; DB 2; Length 16;  
Best Local Similarity 35.7%; Pred. No. 1.5e+03;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NTSQYDITYCFNASE 15  
| | | | | : :  
DB 2 NHSPVRNYCLHAES 15

## RESULT 7

PH1317  
Ig heavy chain DJ region (clone C527-112) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: PH1317  
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.  
J. Exp. Med. 176, 1577-1581, 1992  
A:Title: Predominance of fetal type DJH joining in young children with B precursor ly  
A:Reference number: PH1302; MUID:93094761  
A:Accession: PH1317  
A:Molecule type: DNA  
A:Residues: 1-16 <WAS>  
C:Keywords: heterotetramer; immunoglobulin

Query Match 19.3%; Score 26; DB 2; Length 16;  
Best Local Similarity 44.4%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 TYCFNASAP 16  
| | : : | |  
DB 3 TSCYGGSTP 11

## RESULT 8

S08605  
hypothetical protein 1 estrogen receptor 5'-region - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 18-Aug-2000  
C:Accession: S08605  
R:Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chambon, P.  
EMBO J. 5, 891-897, 1986  
A:Title: The chicken oestrogen receptor sequence: homology with v-erbA and the human  
A:Reference number: S07192; MUID:86247578  
A:Accession: S08605  
A:Status: translation not shown

A:Molecule type: mRNA  
A:Residues: 1-20 <KRU>

A:Cross-references: EMBL:X03805; NID:g63378; PIDN:CAA27431.1; PID:g63379

A:Note: the authors translated the codon TTT for residue 5 as Gly and TTC for residue 16  
C:Superfamily: unassigned leader peptides

Query Match 18.9%; Score 25.5; DB 2; Length 20;

Best Local Similarity 60.0%; Pred. No. 2.2e+03;

Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 10 CFN-ASAPPE 18

Db 4 CFNMISAEPO 13

RESULT 9

PH1458

T-cell receptor beta chain (clone 332/1K) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 17-Mar-1999

C:Accession: PH1458; S26543

R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; K

J. Exp. Med. 177, 811-820, 1993

A:Title: T cell receptor selection by and recognition of two class I major histocompatib

A:Reference number: PH1430; MUID:93171821

A:Accession: PH1458

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Experimental source: cytolytic T-lymphocyte, clone 332/1K

R:Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wid

J. Exp. Med. 176, 439-447, 1992

A:Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor

A:Reference number: S26512; MUID:92364546

A:Accession: S26543

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-12 <CA2>

A:Cross-references: EMBL:X67993

A:Experimental source: cytolytic T-lymphocyte, clone Cw3/C44

C:Superfamily: immunoglobulin homology

C:Keywords: receptor; T-cell

Query Match 18.5%; Score 25; DB 2; Length 12;

Best Local Similarity 45.5%; Pred. No. 1.6e+03;

Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CNTSQDYTYCF 11

Db 1 CASSQGSDYTF 11

RESULT 10

B45474

thrombospondin 2 - bovine (fragment)

N:Alternate names: corticotropin-induced secreted protein (CISP); thrombospondin homolog

C:Species: Bos primigenius taurus (cattle)

C>Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 28-Feb-1997

C:Accession: B45474

R:Pellerin, S.; Lafeuillade, B.; Scherrer, N.; Gagnon, J.; Shi, D.L.; Chambaz, E.M.; Fei

J. Biol. Chem. 268, 4304-4310, 1993

A:Title: Corticotropin-induced secreted protein, an ACTH-induced protein secreted by ad

A:Reference number: A45474; MUID:93179438

A:Accession: B45474

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-15 <PEL>

A:Experimental source: adrenocortical cells

A:Note: sequence extracted from NCBI backbone (NCBIP:125842)

Query Match 18.5%; Score 25; DB 2; Length 15;

Best Local Similarity 57.1%; Pred. No. 1.9e+03;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NTSQYDT 8

Db 3 NQAQFDT 9

RESULT 11

A60839

neurokinin A homolog - marbled electric ray

N:Alternate names: des-Ser(1),pro(2) scyllorhinin II

C:Species: Torpedo marmorata (marbled electric ray)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

C:Accession: A60839

R:Conlon, J.M.; Thim, L.

Gen. Comp. Endocrinol. 71, 383-388, 1988

A:Title: Isolation of the tachykinin, Des(Ser(1))Pro(2)] scyllorhinin II from the inte

A:Reference number: A60839; MUID:89053024

A:Accession: A60839

A:Molecule type: protein

A:Residues: 1-16 <CON>

C:Keywords: amidated carboxyl end; neuropeptide; tachykinin

F:16/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 18.5%; Score 25; DB 2; Length 16;

Best Local Similarity 40.0%; Pred. No. 2.1e+03;

Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 12 NASAPPEDC 21

Db 2 NSKCPDGPDC 11

RESULT 12

PH0763

T-cell receptor beta chain (F15) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999

C:Accession: PH0763

R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A:Title: T cell receptor genes in a series of class I major histocompatibility comple

allelic exclusion and antigen-specific repertoire.

A:Reference number: PH0746; MUID:92078846

A:Accession: PH0763

A:Molecule type: mRNA

A:Residues: 1-16 <CAS>

A:Cross-references: EMBL:X60857; NID:g50933; PIDN:CAA43247.1; PID:g50934

A:Experimental source: T lymphocyte

C:Keywords: T-cell receptor

Query Match 18.5%; Score 25; DB 2; Length 16;

Best Local Similarity 66.7%; Pred. No. 2.1e+03;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 ASAPPE 18

Db 2 ASSPPQ 7

RESULT 13

PH0759

T-cell receptor beta chain (QB7.3.2) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999

C:Accession: PH0759

R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A:Title: T cell receptor genes in a series of class I major histocompatibility comple

allelic exclusion and antigen-specific repertoire.



Job time: 109 sec

A:Reference number: PH0746; MUID:92078846  
A:Accession: PH0759  
A:Molecule type: mRNA  
A:Residues: 1-16 <CAS>  
A:Cross-references: EMBL:X60854; NID:g53878; PIDN:CAA43244.1; PID:g53879  
A:Experimental source: T lymphocyte  
C:Keywords: T-cell receptor

Query Match 18.5%; Score 25; DB 2; Length 16;  
Best Local Similarity 66.7%; Pred. No. 2.1e+03;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 ASAPPE 18  
\* ||:||  
Db 2 ASSPPQ 7

RESULT 14  
A42920  
fatty acid ethyl ester synthase-II - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995  
C:Accession: A42920  
R:Bora, P.S.; Wu, X.; Spilburg, C.A.; Lange, L.G.  
J. Biol. Chem. 267, 13217-13221, 1992  
A:Title: Purification and characterization of fatty acid ethyl ester synthase-II from hu  
A:Reference number: A42920; MUID:92317032  
A:Accession: A42920  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-17 <BOR>  
A:Experimental source: myocardium  
A>Note: sequence extracted from NCBI backbone (NCBIP:107742)

Query Match 18.5%; Score 25; DB 2; Length 17;  
Best Local Similarity 57.1%; Pred. No. 2.2e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 PPEEDCT 22  
||: ||  
Db 3 PPDPDTT 9

RESULT 15  
B24867  
scyllorhinin II - smaller spotted catshark  
C:Species: Scyllorhinus canicula (smaller spotted catshark, smaller spotted dogfish)  
C:Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 08-Dec-1995  
C:Accession: B24867  
R:Conlon, J.N.; Deacon, C.F.; O'Toole, L.; Thim, L.  
FEBS Lett. 200, 111-116, 1986  
A:Title: Scyllorhinin I and II: two novel tachykinins from dogfish gut.  
A:Reference number: A91359; MUID:86192829  
A:Accession: B24867  
A:Molecule type: protein  
A:Residues: 1-18 <CON>  
C:Keywords: amidated carboxyl end; neuropeptide  
F:18/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 18.5%; Score 25; DB 2; Length 18;  
Best Local Similarity 40.0%; Pred. No. 2.3e+03;  
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 12 NASAPPEDC 21  
||: ||  
Db 4 NSKCPDGPDC 13

Search completed: February 21, 2002, 16:31:13



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:30:55 ; Search time 10.02 seconds  
(without alignments)  
84.161 Million cell updates/sec

Title: US-08-753-851-2

Perfect score: 135

Sequence: 1 CNTSQDYTYCFNASAPPEEDCTS 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1274

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	19.3	13	1 CXAL_CONST	P15471 conus stria
2	26	19.3	22	1 LANN_STRMU	P80666 streptococc
3	26	19.3	22	1 SETB_SALTY	P33027 salmonella
4	25	18.5	18	1 TRN2_SCYCA	P08609 scyllorhinu
5	25	18.5	20	1 COXM_THUOB	P80981 thunnus obe
6	25	18.5	20	1 COXM_THUOB	P80980 thunnus obe
7	24.5	18.1	15	1 UC19_MAIZE	P80525 zea mays (m
8	24	17.8	13	1 CXAA_CONST	P28878 conus stria
9	24	17.8	13	1 CXAA_CONST	P28879 conus stria
10	23.5	17.4	16	1 CXAL_CONEP	P56638 conus episc
11	23	17.0	10	1 UH05_RAT	P56573 rattus norv
12	23	17.0	15	1 UC06_MAIZE	P80612 zea mays (m
13	23	17.0	17	1 A45K_MYCBO	P80069 mycobacteri
14	23	17.0	17	1 LCK_RAT	Q01821 rattus norv
15	23	17.0	18	1 GOME_ACAGO	P82358 acanthoscur
16	23	17.0	19	1 DHAB_COMTE	P80704 comamonas t
17	22.5	16.7	23	1 RT21_BOVIN	P82920 bos taurus
18	22	16.3	12	1 UR2B_CATCO	P04559 catostomus
19	22	16.3	12	1 UR2B_CYPCA	P04561 cyprinus ca
20	22	16.3	14	1 SCR3_LEIQU	P45661 leiurus qui
21	22	16.3	20	1 F1BB_ELEMA	P14538 elephas max
22	21	15.6	12	1 TRN2_KASMA	P08614 kassina mac
23	21	15.6	13	1 UHAI_HUMAN	P40928 homo sapien
24	21	15.6	14	1 ECDC_LYMDI	P80940 lymantria d
25	21	15.6	16	1 IBP4_PIG	P24854 sus scrofa
26	21	15.6	19	1 H3_NARPS	P80553 narcissus p
27	21	15.6	20	1 R1PX_CUCPE	P80750 cucurbita p
28	20	14.8	9	1 TAL3_PICJA	P17441 pichia jadi
29	20	14.8	11	1 LSKP_PERAM	P36885 periplaneta
30	20	14.8	14	1 NSK2_SARBU	P41493 sarcophaga
31	20	14.8	15	1 LEC1_PSOSC	P22582 psophocarpu
32	20	14.8	15	1 LEC2_PSOSC	P22585 psophocarpu
33	20	14.8	15	1 UE15_HORVU	P34938 hordeum vul

#### ALIGNMENTS

```

RESULT 1
CXAL_CONST STANDARD; PRT; 13 AA.
AC P15471;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE ALPHA-CONOTOXIN SI (SI)
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6493;
RN [1]
RP SEQUENCE
RX MEDLINE=89062448; PubMed=3196703;
RA Zafaralla G.C., Ramilo C., Gray W.R., Karlstroem R., Olivera B.M.,
RA Cruz L.J.;
RT "Phylogenetic specificity of cholinergic ligands: alpha-conotoxin
SI";
RL Biochemistry 27:7102-7105(1988).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC PIR: A28953; A28953.
DR HSSP; P01519; INOT.
DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13
FT SEQUENCE 13 AA; 1359 MW; DEE90F8EF2457EBD CRC64;
AMIDATION.

Query Match 19.3%; Score 26; DB 1; Length 13;
Best Local Similarity 33.3%; Pred. No. 3.3e+02;
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 10 CFNASAPPEEDC 21
| | | | |
Db 2 CCNPACGPKVSC 13

RESULT 2
LANN_STRMU STANDARD; PRT; 22 AA.
AC P80666;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LANTIBIOTIC MUTACIN B-NY266.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE.

```

```

RC STRAIN=NY266;
RX MEDLINE=97379322; PubMed=9237644;
RA Mota-Meira M., Lacroix C., Lapointe G., Lavoie M.C.;
RT "Purification and structure of mutacin B-Ny266: a new lantibiotic
   produced by Streptococcus mutans.";
RL FEBS Lett. 410:275-279(1997)
CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
   ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
   LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
   CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
   TRANSMEMBRANE PORES.
CC -1- MASS SPECTROMETRY: MW=2270.29; MW_ERR=0.21; METHOD=ELECTROSPRAY.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
DR InterPro: IPR001049; Gallidermin.
DR Pfam: PF02052; Gallidermin; 1.
DR PRINTS: PR00323; GALLIDERMIN.
KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.
FT MOD_RES 5 5 DHA (2,3-DIDEHYDROALANINE).
FT MOD_RES 8 8 D-ABU (AMINOBUTYRIC ACID).
FT MOD_RES 14 14 DHB (2,3-DIDEHYDROBUTYRINE).
FT MOD_RES 19 19 DHA (2,3-DIDEHYDROALANINE).
FT THIOETH 3 7 ALA-S-ALA (LANTHIONINE).
FT THIOETH 8 11 ALA-S-ALA (BETA-METHYLLANTHIONINE).
FT THIOETH 16 21 ALA-S-ALA (LANTHIONINE).
FT THIOETH 19 22 DHA-S-ALA (AVI).
SQ SEQUENCE 22 AA; 2425 MW; 961C1480401F92CE CRC64;

Query Match 19.3%; Score 26; DB 1; Length 22;
Best Local Similarity 37.5%; Pred. No. 5.6e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 TSOVDTYC 10
DB 14 TGSFNSYC 21

RESULT 3
SETB_SALTY STANDARD; PRT; 22 AA.
AC F33027;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SUGAR EFFLUX TRANSPORTER B (FRAGMENT).
GN SETB.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89313694; PubMed=2546043;
RA Geerse R.H., Izso F., Postma P.W.;
RT "The PEP: fructose phosphotransferase system in Salmonella
   typhimurium: Pfr combines enzyme IIfu and pseudo-HPr activities.";
RL Mol. Gen. Genet. 216:517-525(1989).
CC -1- FUNCTION: INVOLVED IN THE EFFLUX OF SUGARS. THE PHYSIOLOGICAL ROLE
   MAY BE THE DETOXIFICATION OF NON-METABOLIZABLE SUGAR ANALOGS. CAN
   TRANSPORT LACTOSE AND GLUCOSE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
   (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE SET FAMILY OF TRANSPORTERS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14243; NOT_ANNOTATED_CDS.

```

```

DR StyGene; SG10423; setB.
KW Transport; Sugar transport; Transmembrane; Inner membrane.
FT TRANSMEM 13 >22 POTENTIAL.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2291 MW; 2849C330172CC5C2 CRC64;

Query Match 19.3%; Score 26; DB 1; Length 22;
Best Local Similarity 54.5%; Pred. No. 5.6e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 ASAPPEEDCTS 23
DB 7 AAAPKSFDLTS 17

RESULT 4
TKN2_SCYCA STANDARD; PRT; 18 AA.
AC P08609;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SCYLIORHININ II (RECTIN).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=86192829; PubMed=2422058;
RA Conlon J.M., Deacon C.F., O'Toole L., Thim L.;
RT "Scyliorhinin I and II: two novel tachykinins from dogfish gut.";
RL FEBS Lett. 200:111-116(1986).
RN [2]
RP SEQUENCE.
RC TISSUE=Small intestine;
RX MEDLINE=95335921; PubMed=7541963;
RA Anderson W.G., Conlon J.M., Hazon N.;
RT "Characterization of the endogenous intestinal peptide that
   stimulates the rectal gland of Scyliorhinus canicula.";
RL Am. J. Physiol. 268:R1359-R1364(1995).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
   EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
   SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
   MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR PIR: B24867; B24867.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT DISULFID 7 13
FT MOD_RES 18 18 AMIDATION.
SQ SEQUENCE 18 AA; 1854 MW; FCCA3FB01E2F3907 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 18;
Best Local Similarity 40.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 12 NASAPPEEDC 21
DB 4 NSKCPDGPDC 13

RESULT 5
COXM_THUOB STANDARD; PRT; 20 AA.
ID COXM_THUOB
AC P80981;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

KL	THEOR. APPL. GENET.	93:997-1005(1996).	
LC	-1-	MISCELLANEOUS: ON THE 2-D GEL-THE DETERMINED PI OF THIS UNKNOWN	
CC	PROTEIN IS:	5.6, ITS MW IS: 18.4 KDA.	
CC	Maize-2DPAGE:	P80625; COLEOPTILE.	
DR	MaizeGDB:	123951; -.	
FT	NON_TER	1	
FT	NON_TER	15	
SO	SEQUENCE	15 AA: 1672 MW: 1CF69D4DA8737F9D CRC64:	

CC	INHIBIT THEM.
DR	PIR; A40312; NTKNAS.
KW	Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW	Venom.
FT	DISULFID 2 7

FT DISULFID 3 13  
 FT MOD\_RES 13 13  
 SQ SEQUENCE 13 AA; 1461 MW; DEF1931982457EBD CRC64;

Query Match 17.8%; Score 24; DB 1; Length 13;  
 Best Local Similarity 30.8%; Pred. No. 6.4e+02;  
 Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 9 YCFNASAPPEDC 21  
 || : : ||  
 Db 1 YCCHPACGKNFDC 13

RESULT 9

CXA2\_CONST STANDARD; PRT; 19 AA.  
 AC P28879;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE ALPHA-CONOTOXIN SII (S2).  
 OS Conus striatus (Striated cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=6493;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Venom;  
 RX MEDLINE=93003172; PubMed=1390774;  
 RA Ramilo C., Zafaralla G.C., Nadasdi L., Hammerland L.G., Yoshikami D.,  
 RA Gray W.R., Kristipati R., Ramchandran J., Miljanich G., Olivera B.M.,  
 RA Cruz L.J.;  
 RT "Novel alpha- and omega-conotoxins from Conus striatus venom.";  
 RL Biochemistry 31:9919-9926(1992).  
 CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
 CC INHIBIT THEM.  
 CC PIR: A44379; A44379.  
 DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Venom.  
 KW DISULFID 2 18  
 FT DISULFID 3 8  
 FT DISULFID 4 14  
 SQ SEQUENCE 19 AA; 1797 MW; A158B024273EF4CF CRC64;

Query Match 17.8%; Score 24; DB 1; Length 19;  
 Best Local Similarity 33.3%; Pred. No. 9.6e+02;  
 Matches 4; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 10 CFNASAPPEDC 21  
 || : : ||  
 Db 3 CCNPACGPNYGC 14

RESULT 10

CXA1\_CONEP STANDARD; PRT; 16 AA.  
 AC P56638;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ALPHA-CONOTOXIN EPI.  
 OS Conus episcopatus (Bishop's cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=88764;  
 RN [1]  
 RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).  
 RX MEDLINE=98376423; PubMed=9708977;  
 RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,  
 RA Alewood P.F., Lewis R.J., Martin J.L.;  
 RT "The 1.1-A resolution crystal structure of [Tyr15]EPI, a novel

RT alpha-conotoxin from Conus episcopatus, solved by direct methods.";  
 RL Biochemistry 37:11425-11433(1998).  
 CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
 CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC  
 CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-  
 CC 3/BETA-4 SUBUNITS.  
 CC PDB; LAOM; 13-JAN-99.  
 DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
 KW Venom; Sulfation; 3D-structure.  
 FT DISULFID 2 8  
 FT DISULFID 3 16  
 FT MOD\_RES 15 15  
 FT MOD\_RES 16 16  
 SQ SEQUENCE 16 AA; 1792 MW; C63385F376C99B4C CRC64;

Query Match 17.4%; Score 23.5; DB 1; Length 16;  
 Best Local Similarity 50.0%; Pred. No. 9.5e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 CNTSQYDTYC 10  
 || : ||  
 Db 8 CMNPNPD-YC 16

RESULT 11

UH05\_RAT STANDARD; PRT; 10 AA.  
 AC P56573;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF HEART TISSUE (SPOT P5) (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=WISTAR; TISSUE=Heart;  
 RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,  
 RA Jungblut P.R.;  
 RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 8.3, ITS MW IS: 30 KDA.  
 CC UNSURE 9 9  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1269 MW; 00CBA4BB46C5BABB CRC64;

Query Match 17.0%; Score 23; DB 1; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 6.9e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 SQYDTY 9  
 |||| :  
 Db 5 SQYDGF 10

RESULT 12

UC06\_MAIZE STANDARD; PRT; 15 AA.  
 ID UC06\_MAIZE  
 AC P80612;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 1131)  
 DE (FRAGMENT).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.

```

OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.8, ITS MW IS: 71.0 KDA.
DR Maize-2DPAGE; P80612; COLEOPTILE.
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 1390 MW; 7005E22830F23D61 CRC64;

Query Match 17.0%; Score 23; DB 1; Length 15;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 ASAPP 17
Db 2 AAAPP 6

RESULT 13
A45K_MYCBO
ID A45K_MYCBO STANDARD; PRT; 17 AA.
AC P80069;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 45/47 KDA ANTIGEN (FRAGMENT).
OS Mycobacterium bovis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE.
RC STRAIN=BCG / PARIS 1173 P2;
RX MEDLINE=93138802; PubMed=8423100;
RA Romain F., Laqueyrie A., Millitzer P., Pescher P., Chavarot P.,
RA Lagranderie M., Auregan G., Gheorghiu M., Marchal G.A.;
RT "Identification of a Mycobacterium bovis BCG 45/47-kilodalton antigen
RT complex, an immunodominant target for antibody response after
RT immunization with living bacteria.";
RL Infect. Immun. 61:742-750(1993).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: TO M.LEPRAE NL43, AND M.TUBERCULOSIS MPT32.
DR PIR: A49237; A49237.
KW Antigen.
FT NON_TER 17
SQ SEQUENCE 17 AA; 1521 MW; 4492CC389D9D9893 CRC64;

Query Match 17.0%; Score 23; DB 1; Length 17;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 ASAPP 17
Db 12 AAAPP 16

RESULT 14
LCK_RAT
ID LCK_RAT STANDARD; PRT; 17 AA.
AC Q01621;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

---

```

DE DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE LCK (EC 2.7.1.112) (P56-LCK)
DE (FRAGMENT).
GN LCK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER;
RX MEDLINE=931411260; PubMed=8423992;
RA Shin S., Steffen D.L.;
RT "Frequent activation of the lck gene by promoter insertion and
RT aberrant splicing in murine leukemia virus-induced rat lymphomas.";
RL Oncogene 8:141-149(1993).
CC -!- FUNCTION: MAY PARTICIPATE IN ANTIGEN-INDUCED T-CELL ACTIVATION.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER
CC CD4 OR CD8.
CC -!- DISEASE: PROVIRAL INSERTION UPSTREAM OF THE LCK GENE CAUSES
CC OVEREXPRESSION, LEADING TO THE DEVELOPMENT OF THYMIC LYMPHOMA.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE SRC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z15029; CAA78748.1;
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR001245; Tyr_kin.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; PARTIAL.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; PARTIAL.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; PARTIAL.
KW Proto-oncogene; Tyrosine-protein kinase; Phosphorylation; Transferase;
KW ATP-binding; Myristate; Palmitate; Lipoprotein.
FT INIT_MET 0 0 PROBABLE.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT LIPID 2 2 PALMITATE (BY SIMILARITY).
FT LIPID 4 4 PALMITATE (BY SIMILARITY).
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 1914 MW; 29B848FB8F0EB776 CRC64;

Query Match 17.0%; Score 23; DB 1; Length 17;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 SAPPEED 20
Db 5 SSNPEDD 11

RESULT 15
GOME_ACAGO
ID GOME_ACAGO STANDARD; PRT; 18 AA.
AC P82358;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GOMESIN.
OS Acanthoscurria gomesiana.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Mygalomorphae; Theraphosidae; Acanthoscurria.
OX NCBI_TaxID=115339;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Hemocyte;

```

RX MEDLINE-20517902; PubMed-10942757;  
RA Silva P.I. Jr., Daffre S., Bulet P.;  
RT "Isolation and characterization of gomesin, an 18-residue cysteine-  
RT rich defense peptide from the spider Acanthoscurria gomesiana  
RT hemocytes with sequence similarities to horseshoe crab antimicrobial  
RT peptides of the tachyplesin family.";  
RL J. Biol. Chem. 275:33464-33470(2000).  
CC -!- FUNCTION: ACTIVE AGAINST SEVERAL GRAM-POSITIVE BACTERIA SUCH AS  
CC BACILLUS SPP, STAPHYLOCOCCUS SPP AND E.FAECALIS, SEVERAL GRAM-  
CC NEGATIVE BACTERIA SUCH AS E.COLI, K.PNEUMONIAE, P.AERUGINOSA AND  
CC SALMONELLA SPP, FILAMENTOUS FUNGI SUCH AS N.CRASSA, T.VIRIDAE  
CC AND YEASTS SUCH AS C.ALBICANS. IT IS ACTIVE AGAINST THE PARASITE  
CC L.AMAZONENSIS AS WELL. IT SHOWS HEMOLYTIC ACTIVITY.  
CC -!- TISSUE SPECIFICITY: HEMOCYTES.  
CC -!- MASS SPECTROMETRY: MW=2270.4; METHOD=WALDI.  
KW Amidation; Antibiotic; Fungicide; Hemolysis.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 18 18 AMIDATION.  
FT DISULFID 2 15  
FT DISULFID 6 11  
SQ SEQUENCE 18 AA; 2293 MW; 25ED499784908913 CRC64;

Query Match 17.0%; Score 23; DB 1; Length 18;  
Best Local Similarity 40.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CNTSQYDTC 10  
| : |||  
DB 6 CYKQRCVTC 15

Search completed: February 21, 2002, 16:33:57  
Job time: 182 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:30:39 ; Search time 22.06 seconds  
(without alignments)  
152.505 Million cell updates/sec

Title: US-08-753-851-2

Perfect score: 135

Sequence: 1 CNTSQDYTYCFNASAPPEEDCTS 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 6766

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP TREMBL.17.\*  
2: SP Archaea.\*  
3: SP Bacteria.\*  
4: SP Fungi.\*  
5: SP Human.\*  
6: SP Invertebrate.\*  
7: SP Mammal.\*  
8: SP MHC.\*  
9: SP Organelle.\*  
10: SP Phage.\*  
11: SP Plant.\*  
12: SP Rodent.\*  
13: SP Virus.\*  
14: SP Vertebrate.\*  
15: SP Unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34.5	25.6	22	12 Q85481	Q85481 rous sarcoma
2	34	25.2	23	2 Q9X7J5	Q9X7J5 pseudomonas
3	33	24.4	22	2 Q9R5U3	Q9R5U3 pseudomonas
4	29	21.5	18	2 Q9R5U2	Q9R5U2 pseudomonas
5	28	20.7	11	5 Q9292	Q9292 drosophila
6	28	20.7	20	11 Q9QV27	Q9QV27 rattus sp.
7	27.5	20.4	20	5 Q46158	Q46158 lumbricus r
8	27	20.0	16	11 Q9QUY8	Q9QUY8 cricetus
9	27	20.0	19	2 Q9R5G6	Q9R5G6 fusobacteri
10	27	20.0	19	11 Q9QVB8	Q9QVB8 mus sp. tru
11	27	20.0	19	12 Q9I329	Q9I329 human immun
12	27	20.0	19	12 Q9WJ51	Q9WJ51 human immun
13	26.5	19.6	22	12 Q9PXA8	Q9PXA8 human immun
14	26	19.3	13	4 Q9UCU2	Q9UCU2 homo sapien
15	26	19.3	14	2 Q44172	Q44172 synechococ
16	26	19.3	18	13 P82110	P82110 rana pipien
17	26	19.3	20	4 Q9UMI8	Q9UMI8 homo sapien
18	26	19.3	20	5 Q9BP45	Q9BP45 conus penna
19	25	18.5	13	4 Q14890	Q14890 homo sapien

20 25 18.5 15 10 Q9S914  
21 25 18.5 16 2 Q9R963  
22 25 18.5 16 2 Q53399  
23 25 18.5 17 4 Q9UD06  
24 25 18.5 18 10 Q9S915  
25 25 18.5 19 4 Q9BYK5  
26 24 17.8 8 6 Q9BFC3  
27 24 17.8 8 6 Q9BFC2  
28 24 17.8 8 6 Q9BFC1  
29 24 17.8 8 6 Q9BFC0  
30 24 17.8 8 6 Q9BFB9  
31 24 17.8 8 6 Q9BFB8  
32 24 17.8 8 6 Q9BFB7  
33 24 17.8 8 6 Q9BFB6  
34 24 17.8 8 6 Q9BFB5  
35 24 17.8 8 6 Q9BFB4  
36 24 17.8 8 6 Q9BFB3  
37 24 17.8 8 6 Q9BFB2  
38 24 17.8 8 6 Q9BFB1  
39 24 17.8 8 6 Q9BFB0  
40 24 17.8 8 6 Q9BFA9  
41 24 17.8 8 6 Q9BFA8  
42 24 17.8 8 6 Q9BFA6  
43 24 17.8 8 6 Q9BFA5  
44 24 17.8 8 6 Q9BFA4  
45 24 17.8 8 6 Q9BFA3

Q9S914 hordeum vul  
Q9R963 helicobacte  
Q53399 bacillus th  
Q9UD06 homo sapien  
Q9S915 triticum tu  
Q9BYK5 homo sapien  
Q9BFC3 didelphis m  
Q9BFC2 macropus eu  
Q9BFC1 choleopus h  
Q9BFC0 choleopus d  
Q9BFB9 euphractus  
Q9BFB8 chaetophrac  
Q9BFB7 tamandua te  
Q9BFB6 myrmecophag  
Q9BFB5 erinaceus c  
Q9BFB4 talpa alta  
Q9BFB3 condylura c  
Q9BFB2 sorex arane  
Q9BFB1 echinops te  
Q9BFB0 trichechus  
Q9BFA9 procavia ca  
Q9BFA8 loxodonta a  
Q9BFA6 orycteropus  
Q9BFA5 cynocephalu  
Q9BFA4 tupaia mino  
Q9BFA3 lemur catta

#### ALIGNMENTS

RESULT 1  
Q85481 PRELIMINARY; PRT; 22 AA.  
AC Q85481;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DE 01-NOV-1996 (TREMREL. 01, Last annotation update)  
DE P19GAG AND P2GAG PROTEINS (FRAGMENT).  
GN P2GAG.  
OS Rous sarcoma virus.  
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.  
OX NCBI\_TaxID=11886;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TDPH2010;  
RA Hara H., Tanabe T., Kaji A.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; M77660; AAA91044.1; -  
FT NON\_TER 1 1  
SQ SEQUENCE 22 AA; 2087 MW; 6E160516483D4435 CRC64;

Query Match 25.6%; Score 34.5; DB 12; Length 22;  
Best Local Similarity 47.1%; Pred. No. 99;  
Matches 8; Conservative 0; Mismatches 2; Indels 7; Gaps 1;

Qy 1 CNTSQDYTYCFNASAPP 17  
|||  
Db 8 CN-----CATASAPP 17

RESULT 2  
Q9X7J5 PRELIMINARY; PRT; 23 AA.  
AC Q9X7J5;  
DT 01-NOV-1999 (TREMREL. 12, Created)  
DT 01-NOV-1999 (TREMREL. 12, Last sequence update)  
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)  
DE HYPOTHETICAL 2.6 KDA PROTEIN (FRAGMENT).  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
Pseudomonas.

```

OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZOBELL, ATCC14405;
RX MEDLINE=9226217; PubMed=10209742;
RA Vollack K.U., Hartig E., Korner H., Zumft W.G.;
RT "Multiple transcription factors of the FNR family in denitrifying
RT Pseudomonas stutzeri: characterization of four fnr-like genes,
RT regulatory responses and cognate metabolic processes.";
RL Mol. Microbiol. 31:1681-1694(1999).
DR EMBL: AJ131715; CAB40904.1;
KW Hypothetical protein.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2618 MW; EF68EA0F803A5EBD CRC64;

Query Match 25.2%; Score 34; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 9 YCFNASAPPEEDCTS 23
   |||::|::|::|
Db 6 YCOQDNPQAECSN 20

RESULT 3
Q9R5U3 PRELIMINARY; PRT; 22 AA.
AC Q9R5U3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE 2-AMINO-BENZOATE-COA LIGASE-E2
OS Pseudomonas (fluorescent pseudomonads).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
OX NCBI_TaxID=286;
RN [1]
RP SEQUENCE.
RX MEDLINE=91358327; PubMed=1885526;
RA Altnschmidt U., Oswald B., Fuchs G.;
RT "Purification and characterization of benzoate-coenzyme A ligase and
RT 2-aminobenzoate-coenzyme A ligases from a denitrifying Pseudomonas
RT sp.";
RL J. Bacteriol. 173:5494-5501(1991).
SQ SEQUENCE 22 AA; 2526 MW; 5D37FA76363F8792 CRC64;

Query Match 24.4%; Score 33; DB 2; Length 22;
Best Local Similarity 33.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 3 TSQYDTYCFNASAPPEEDCTS 23
   |||::|::|::|
Db 1 TSHVDTFARDXPPTTEQOTES 21

RESULT 4
Q9R5U2 PRELIMINARY; PRT; 18 AA.
AC Q9R5U2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE 2-AMINO-BENZOATE-COA LIGASE-E3
OS Pseudomonas (fluorescent pseudomonads).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
OX NCBI_TaxID=286;
RN [1]
RP SEQUENCE.
RX MEDLINE=91358327; PubMed=1885526;
RA Altnschmidt U., Oswald B., Fuchs G.;
RT "Purification and characterization of benzoate-coenzyme A ligase and
RT 2-aminobenzoate-coenzyme A ligases from a denitrifying Pseudomonas

```

```

RT sp.";
RL J. Bacteriol. 173:5494-5501(1991).
SQ SEQUENCE 18 AA; 2061 MW; C43F8799692771EF CRC64;

Query Match 21.5%; Score 29; DB 2; Length 18;
Best Local Similarity 35.3%; Pred. No. 6e+02;
Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 TSQYDTYCFNASAPPEE 19
   |||::|::|::|
Db 1 TSHVDTFARDXLPPXEQ 17

RESULT 5
Q9R292 PRELIMINARY; PRT; 11 AA.
AC Q9R292;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE BICOID PROTEIN (FRAGMENT).
GN BCD.
OS Drosophila heteroneura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=32382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91184004; PubMed=2081457;
RA MacDonald P.M.;
RT "bicoid mRNA localization signal: phylogenetic conservation of
RT function and RNA secondary structure.";
RL Development 110:161-171(1990).
CC -1- FUNCTION: BICOID IS SEGMENT-POLARITY PROTEIN THAT PROVIDES
CC POSITIONAL CUES FOR THE DEVELOPMENT OF HEAD AND THORACIC SEGMENTS.
CC BCD REGULATES THE EXPRESSION OF ZYGOTIC GENES, POSSIBLY THROUGH
CC ITS HOMEODOMAIN, AND INHIBITS THE ACTIVITY OF OTHER MATERNAL GENE
CC PRODUCTS. IT IS POSSIBLE THAT BCD ALSO BINDS RNA.
DR EMBL: M32125; AAA28386.1;
DR FlyBase; FBgn0012352; Dhet\bcd.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Segmentation polarity protein; Transcription regulation; RNA-binding.
FT NON_TER 1
SQ SEQUENCE 11 AA; 1221 MW; 8CE802305DD9D6C1 CRC64;

Query Match 20.7%; Score 28; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 YCFN 12
   ||||
Db 8 YCFN 11

RESULT 6
Q9QV27 PRELIMINARY; PRT; 20 AA.
AC Q9QV27;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE INTEGRIN BETA 1 CHAIN.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=94291763; PubMed=8020590;
RA Forsberg E., Ek B., Engstrom A., Johansson S.;

```

RT. "Purification and characterization of integrin alpha 9 beta 1.";  
 RL Exp. Cell Res. 213:183-190(1994).  
 SQ SEQUENCE 20 AA; 2335 MW; 50220366111FB2D8 CRC64;

Query Match 20.7%; Score 28; DB 11; Length 20;  
 Best Local Similarity 44.4%; Pred. No. 9.6e+02;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NTSQYDTYC 10  
 I: | | |  
 Db 10 NSEDMDAYC 18

RESULT 7  
 ID O46158 PRELIMINARY; PRT; 20 AA.  
 AC O46158;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE 40S RIBOSOMAL PROTEIN S27A (FRAGMENT).  
 OS Lumbricus rubellus (Humus earthworm).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
 OC Lumbricina; Lumbricidae; Lumbricus.  
 OX NCBI\_TaxID=35632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sturzenbaum S.R., Kille P., Morgan J.;  
 RL Submitted (DGC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ223199; CAA11174.1; -;  
 DR InterPro; IPR002906; Ribosomal\_S27.  
 DR Pfam; PF01599; Ribosomal\_S27; 1.  
 KW Ribosomal protein.  
 FT NON\_TER  
 SQ SEQUENCE 20 AA; 2388 MW; 7685DE00FB23B91B CRC64;

Query Match 20.4%; Score 27.5; DB 5; Length 20;  
 Best Local Similarity 58.3%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 8 TYCFNASAPPEE 19  
 I: | | | | |  
 Db 11 TYVFN---KPEE 19

RESULT 8  
 ID Q90UY8 PRELIMINARY; PRT; 16 AA.  
 AC Q90UY8;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE HISTONE H1 (FRAGMENT).  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96070893; PubMed=7499230;  
 RA Gurley L.R., Valdez J.G., Buchanan J.S.;  
 RT "Characterization of the mitotic specific phosphorylation site of histone H1. Absence of a consensus sequence for the p34cdc2/cyclin B kinase.";  
 RL J. Biol. Chem. 270:27653-27660(1995).  
 SQ SEQUENCE 16 AA; 1479 MW; 75EB488737288C8A CRC64;

Query Match 20.0%; Score 27; DB 11; Length 16;  
 Best Local Similarity 71.4%; Pred. No. 1.1e+03;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 13 ASAPPEE 19  
 I: | | | | |  
 Db 9 AAAPPAAE 15

RESULT 9  
 ID Q9R5G6 PRELIMINARY; PRT; 19 AA.  
 AC Q9R5G6;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE 2-HYDROXYGLUTARYL-COA DEHYDRATASE GAMMA-SUBUNIT.  
 OS Fusobacterium nucleatum.  
 OC Bacteria; Fusobacteria; Fusobacterium.  
 OX NCBI\_TaxID=851;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93038007; PubMed=1417419;  
 RA Klees A.G., Linder D., Buckel W.;  
 RT "2-Hydroxyglutaryl-CoA dehydratase from Fusobacterium nucleatum (subsp. nucleatum): an iron-sulfur flavoprotein.";  
 RL Arch. Microbiol. 158:294-301(1992).  
 SQ SEQUENCE 19 AA; 2255 MW; 1E3C577D5B00802D CRC64;

Query Match 20.0%; Score 27; DB 2; Length 19;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NTSQYDTYCF 11  
 I: | | | | |  
 Db 2 NIXEYDDFIF 11

RESULT 10  
 ID Q9QVB8 PRELIMINARY; PRT; 19 AA.  
 AC Q9QVB8;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE TRUE TISSUE KALLIKREIN (EC 3.4.21.35) (FRAGMENT).  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92348355; PubMed=1639762;  
 RA Peters J., Takahashi S., Tada M., Miyake Y.;  
 RT "mGK-6-derived true tissue kallikrein is synthesized, processed, and targeted through a regulated secretory pathway in mouse pituitary AtT-20 cells.";  
 RL J. Biochem. 111:643-648(1992).  
 SQ SEQUENCE 19 AA; 2296 MW; 250436479C0E6626 CRC64;

Query Match 20.0%; Score 27; DB 11; Length 19;  
 Best Local Similarity 45.5%; Pred. No. 1.3e+03;  
 Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 10 CFNASAPPEED 20  
 I: | | | | |  
 Db 9 CVNLKLLPNED 19

RESULT 11  
 ID O91329 PRELIMINARY; PRT; 19 AA.  
 AC O91329;

DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT SI FROM FRANCE;  
 RX MEDLINE=98285741; PubMed=9621043;  
 RA Belec L., Si Mohamed A., Muller-Trutwin M.C., Gilquin J., Gutmann L.,  
 RA Safar M., Barre-Sinoussi F., Kazatchkine M.D.;  
 RT "Genetically related human immunodeficiency virus type 1 in three  
 RT adults of a family with no identified risk factor for intrafamilial  
 RT transmission.";  
 RL J. Virol. 72:5831-5839(1998).  
 DR EMBL; U87220; AAC32980.1; -;  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW Envelope protein.  
 KX  
 FT NON\_TER  
 SQ SEQUENCE 19 AA; 2324 MW; 379CB14A9E073911 CRC64;  
 Query Match 20.0%; Score 27; DB 12; Length 19;  
 Best Local Similarity 42.9%; Pred. No. 1.3e+03;  
 Matches 6; Conservative 3; Mismatches 1; Indels 4; Gaps 1;  
 QY 1 CNTSQDYTYCFNAS 14  
 III: I II::  
 Db 4 CNTTQ----LFNST 13

RESULT 12  
 Q9WJBI  
 ID Q9WJBI PRELIMINARY; PRT; 19 AA.  
 AC Q9WJBI  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT SI FROM FRANCE;  
 RX MEDLINE=98285741; PubMed=9621043;  
 RA Belec L., Si Mohamed A., Muller-Trutwin M.C., Gilquin J., Gutmann L.,  
 RA Safar M., Barre-Sinoussi F., Kazatchkine M.D.;  
 RT "Genetically related human immunodeficiency virus type 1 in three  
 RT adults of a family with no identified risk factor for intrafamilial  
 RT transmission.";  
 RL J. Virol. 72:5831-5839(1998).  
 DR EMBL; U87216; AAC32976.1; -;  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW Envelope protein.  
 KX  
 FT NON\_TER  
 SQ SEQUENCE 19 AA; 2294 MW; 3781714A9E073911 CRC64;

Query Match 20.0%; Score 27; DB 12; Length 19;  
 Best Local Similarity 42.9%; Pred. No. 1.3e+03;  
 Matches 6; Conservative 3; Mismatches 1; Indels 4; Gaps 1;  
 QY 1 CNTSQDYTYCFNAS 14  
 III: I II::  
 Db 4 CNTTQ----LFNST 13

RESULT 13  
 Q9PXA8  
 ID Q9PXA8 PRELIMINARY; PRT; 22 AA.  
 AC Q9PXA8;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE A/ASE PROTEIN (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95100972; PubMed=7802674;  
 RA Sallberg M., Sherefa K., Zhang Z.X.;  
 RT "The antigen/antibody specificity exchanger: a new peptide based tool  
 RT for re-directing antibodies of other specificities to recognize the V3  
 RT domain of HIV-1 gp120.";  
 RL Biochem. Biophys. Res. Commun. 205:1386-1390(1994).  
 SQ SEQUENCE 22 AA; 2704 MW; 36E8F13D871579CE CRC64;

Query Match 19.6%; Score 26.5; DB 12; Length 22;  
 Best Local Similarity 36.8%; Pred. No. 1.8e+03;  
 Matches 7; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 1 CNTSOYD--TYCFNASAP 16  
 I: II I I I I I  
 Db 1 CDLIYYEEDYYFPNPAP 19

RESULT 14  
 Q9UCU2  
 ID Q9UCU2 PRELIMINARY; PRT; 13 AA.  
 AC Q9UCU2;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE INTERLEUKIN-6-23-25 KDA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=91355644; PubMed=1883960;  
 RA May L.T., Shaw J.E., Khanna A.K., Zabriskie J.B., Sehgal P.B.;  
 RT "Marked cell-type-specific differences in glycosylation of human  
 RT interleukin-6.";  
 RL Cytokine 3:204-211(1991).  
 SQ SEQUENCE 13 AA; 1281 MW; 3397986865BFAB18 CRC64;

Query Match 19.3%; Score 26; DB 4; Length 13;  
 Best Local Similarity 62.5%; Pred. No. 1.3e+03;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 ASAPPEED 20  
 I I I I I I I I  
 Db 1 APVPPGED 8

RESULT 15  
 Q44172  
 ID Q44172 PRELIMINARY; PRT; 14 AA.  
 AC Q44172;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE PHYCOCYANIN, ALPHA AND BETA SUBUNITS (FRAGMENT).  
 GN CPCA.  
 OS Synechococcus sp. (strain PCC 7002) (Agmenellum quadruplicatum).

OX Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
RN NCBI\_TaxID=32049;  
RP SEQUENCE FROM N.A.  
RC STRAIN=PR6009;  
RX MEDLINE=90314661; PubMed=2114862;  
RA de Lorimier R., Guglielmi G., Bryant D.A., Stevens S.E.;  
RT "Structure and mutation of a gene encoding a Mr 33,000 phycocyanin-  
associated linker polypeptide";  
RL Arch. Microbiol. 153:541-549(1990).  
DR EMBL; X81868; CAA57456.1; -;  
DR HSP; P07122; ICPC.  
FT NON\_TER 1  
SQ SEQUENCE 14 AA; 1601 MW; 9BD40C7EF60CDA5A CRC64;

Query Match 19.3%; Score 26; DB 2; Length 14;  
Best Local Similarity 45.5%; Pred. No. 1.4e+03;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TSQYDTYCFNA 13  
|: | | | |  
Db 2 TNNYIDYAINA 12

Search completed: February 21, 2002, 16:33:41  
Job time: 182 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:27:39 ; Search time 23.62 seconds

(without alignments)  
72.129 Million cell updates/sec

Title: US-08-753-851-2

Perfect score: 135

Sequence: 1 CMTSQDYTCFNASAPPEEDCTS 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 213999

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1:	/SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135	100.0	23	15 AAR53474	CD44 peptide CD44-
2	45	33.3	11	15 AAR53486	CD44 peptide CD44-
3	43	31.9	13	15 AAR52917	TK-SH2 association
4	41	30.4	13	15 AAR52918	TK-SH2 association
5	41	30.4	13	15 AAR52889	TK-SH2 association
6	41	30.4	13	15 AAR52890	TK-SH2 association
7	40	29.6	18	21 AAB25282	Cytochemical assay
8	38	28.1	14	22 AAG98102	Human SNP associat
9	37	27.4	21	22 AAB48871	Mutant human insul
10	35	25.9	17	21 AAY85630	Snake venom derive
11	34	25.2	13	21 AAY98406	Alpha D peptide de

12	34	25.2	15	22 AAJ03736	Hepatitis C virus
13	34	25.2	15	22 AAJ03739	Hepatitis C virus
14	34	25.2	21	18 AAW42772	Peptide derived fr
15	33	24.4	20	22 AAB82056	HCV antigen, NS3 I
16	33	24.4	22	21 AAB21106	Cow's milk beta-la
17	32	23.7	9	22 AAJ03719	Hepatitis C virus
18	32	23.7	9	22 AAJ03732	Hepatitis C virus
19	32	23.7	14	17 AAW16708	pl85 binding and e
20	32	23.7	14	17 AAW16709	pl85 binding and e
21	32	23.7	14	17 AAW16771	pl85 binding and e
22	32	23.7	14	17 AAW16772	peptide encoded by
23	32	23.7	14	22 AAB57304	pl85 binding and e
24	32	23.7	15	17 AAW16692	pl85 binding and e
25	32	23.7	15	17 AAW16693	pl85 binding and e
26	32	23.7	15	17 AAW16756	pl85 binding and e
27	32	23.7	15	17 AAW16756	peptide derived fr
28	32	23.7	15	22 AAG63556	Cytochemical assay
29	32	23.7	18	21 AAB29279	Tumour necrosis fa
30	32	23.7	18	21 AAB29280	TNF-antagonist pep
31	31.5	23.3	11	20 AAW94628	Osteoclastogenesis
32	31.5	23.3	11	21 AAB17050	GP1b alpha peptide
33	31.5	23.3	11	22 AAB70303	GP1b alpha peptide
34	31.5	23.3	15	12 AAR13137	GP1b-alpha-derived
35	31.5	23.3	15	14 AAR38577	Canine immunoglobi
36	31.5	23.3	20	18 AAW24104	GP1b-alpha-derived
37	31.5	23.3	21	14 AAR38579	Human antithrombin
38	31.5	23.3	21	18 AAW14605	c-raf-related pept
39	31.5	23.3	22	9 AAP81130	Hepatitis C virus
40	31	23.0	8	22 AAJ00095	Hepatitis C virus
41	31	23.0	8	22 AAJ01327	Hepatitis C virus
42	31	23.0	8	22 AAJ02152	Hepatitis C virus
43	31	23.0	9	22 AAJ00888	Hepatitis C virus
44	31	23.0	9	22 AAJ02180	Hepatitis C virus
45	31	23.0	9	22 AAJ02943	Hepatitis C virus

#### ALIGNMENTS

RESULT 1

AA53474

ID AAR53474 standard; peptide; 23 AA.

XX AC AAR53474;

XX AC

DT 01-DEC-1994 (first entry)

XX XX

DE CD44 peptide CD44-2.

XX XX

KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;

KW arthritis; tumor cell metastasis; autoimmune disease;

KW immunosuppressive.

XX XX

OS Homo sapiens.

XX XX

PN WO9409811-A.

XX XX

PD 11-MAY-1994.

XX XX

PF 29-OCT-1993; 93WO-US10412.

XX XX

PR 30-OCT-1992; 92US-0973339.

XX XX

PA (UYDU-) UNIV DUKE.

XX XX

PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;

XX XX

DR WPI; 1994-167121/20.

XX XX

PT Use of CD44 protein and new peptide derivs - for developing prods

PT for inflammation, immune-mediated tissue damage and tumour cell

PT metastasis

XX XX

**II**

AA AAR52918;

XX 07-NOV-1994 (first entry)  
 XX TK-SH2 association inhibitory peptide.  
 XX Tyrosine kinase; SH2 domains; inhibition of association;  
 KW abl; Epidermal Growth Factor Receptor;  
 KW phosphotyrosine residue; control proliferative disease;  
 KW control cancer; TK; EGFR; Src-homology domains.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 4  
 FT /label= other  
 FT /note= "phosphotyrosine"  
 XX WO9407913-A.  
 PN 14-APR-1994.  
 XX 22-SEP-1993; 93WO-US08996.  
 PD 25-SEP-1992; 92US-0951241.  
 PR 15-SEP-1993; 93US-0122028.  
 XX (WARN ) WARNER LAMBERT CO.  
 PA Dobrusin EM, Maclean D, McNamara DJ, Soltiel AR;  
 PI Thieme-sefler, A;  
 PI WPI; 1994-135508/16.  
 DR  
 XX New peptide(s) with a tyrosine auto-phosphorylation site -  
 XX inhibit tyrosine kinase association with regulatory proteins,  
 PT used for treating, e.g., hyper-proliferative, viral, allergic  
 PT inflammatory, auto-immune and cardiovascular diseases  
 XX Table 1; Page 13; 50pp; English.  
 PS The peptide inhibits the association of a tyrosine kinase with its  
 XX cellular substrates and effectively uncouples the tyrosine kinase  
 CC from specific signal transduction pathways. The peptide inhibition  
 CC is probably the result of competition between the phosphorylated  
 CC peptide and the phosphorylated receptor for the same binding site  
 CC on the SH2 domains. The inventors claim the peptides can be used  
 CC for the control of several proliferative diseases, eg. cancer.  
 XX Sequence 13 AA;  
 SQ  
 Query Match 30.4%; Score 41; DB 15; Length 13;  
 Best Local Similarity 63.6%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 9 YCFNASAPPEE 19  
 DB 11:1 111  
 2 ycynpshnpee 12  
 RESULT 5  
 ID AAR52889  
 XX AAR52889 standard; peptide; 13 AA.  
 AC AAR52889;  
 XX 07-NOV-1994 (first entry)  
 DT TK-SH2 association inhibitory peptide.  
 DE Tyrosine kinase; SH2 domains; inhibition of association;  
 XX abl; Epidermal Growth Factor Receptor;  
 KW phosphotyrosine residue; control proliferative disease;  
 KW

KW control cancer; TK; EGFR; Src-homology domains.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FH Modified-site 2  
 FT /label= other  
 FT /note= "phosphotyrosine"  
 XX WO9407913-A.  
 PN 14-APR-1994.  
 XX 22-SEP-1993; 93WO-US08996.  
 PD 25-SEP-1992; 92US-0951241.  
 PR 15-SEP-1993; 93US-0122028.  
 XX (WARN ) WARNER LAMBERT CO.  
 PA Dobrusin EM, Maclean D, McNamara DJ, Soltiel AR;  
 PI Thieme-sefler, A;  
 PI WPI; 1994-135508/16.  
 DR  
 XX New peptide(s) with a tyrosine auto-phosphorylation site -  
 XX inhibit tyrosine kinase association with regulatory proteins,  
 PT used for treating, e.g., hyper-proliferative, viral, allergic  
 PT inflammatory, auto-immune and cardiovascular diseases  
 XX Table 1; Page 11; 50pp; English.  
 PS The peptide inhibits the association of a tyrosine kinase with its  
 XX cellular substrates and effectively uncouples the tyrosine kinase  
 CC from specific signal transduction pathways. The peptide inhibition  
 CC is probably the result of competition between the phosphorylated  
 CC peptide and the phosphorylated receptor for the same binding site  
 CC on the SH2 domains. The inventors claim the peptides can be used  
 CC for the control of several proliferative diseases, eg. cancer.  
 XX Sequence 13 AA;  
 SQ  
 Query Match 30.4%; Score 41; DB 15; Length 13;  
 Best Local Similarity 63.6%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 9 YCFNASAPPEE 19  
 DB 11:1 111  
 2 ycynpshnpee 12  
 RESULT 6  
 ID AAR52890  
 XX AAR52890 standard; peptide; 13 AA.  
 AC AAR52890;  
 XX 07-NOV-1994 (first entry)  
 DT TK-SH2 association inhibitory peptide.  
 DE Tyrosine kinase; SH2 domains; inhibition of association;  
 XX abl; Epidermal Growth Factor Receptor;  
 KW phosphotyrosine residue; control proliferative disease;  
 KW control cancer; TK; EGFR; Src-homology domains.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 4  
 FT /label= other  
 FT /note= "phosphotyrosine"  
 FT



XX WO9407913-A.  
 XX 14-APR-1994.  
 XX 22-SEP-1993; 93WO-US08996.  
 XX 25-SEP-1992; 92US-0951241.  
 PR 15-SEP-1993; 93US-0122028.  
 XX (WARN ) WARNER LAMBERT CO.  
 XX Dobrusin EM, Maclean D, McNamara DJ, Soltiel AR;  
 PI Thieme-sefler, A;  
 XX WPI; 1994-135508/16.  
 XX New peptide(s) with a tyrosine auto-phosphorylation site -  
 PT inhibit tyrosine kinase association with regulatory proteins,  
 PT used for treating, e.g., hyper-proliferative, viral, allergic  
 PT inflammatory, auto-immune and cardiovascular diseases  
 XX Table 1; Page 11; 50pp; English.  
 XX The peptide inhibits the association of a tyrosine kinase with its  
 CC cellular substrates and effectively uncouples the tyrosine kinase  
 CC from specific signal transduction pathways. The peptide inhibition  
 CC is probably the result of competition between the phosphorylated  
 CC peptide and the phosphorylated receptor for the same binding site  
 CC on the SH2 domains. The inventors claim the peptides can be used  
 CC for the control of several proliferative diseases, eg. cancer.  
 XX Sequence 13 AA;  
 SQ  
 Query Match 30.4%; Score 41; DB 15; Length 13;  
 Best Local Similarity 63.6%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 9 YCFNASAPPEE 19  
 DB 2 ycylnpshnpee 12  
 ||:| | |||  
 RESULT 7  
 AAB29282  
 ID AAB29282 standard; Peptide; 18 AA.  
 AC AAB29282;  
 XX 12-FEB-2001 (first entry)  
 DT  
 XX Cytochemical assay quality control 1D5 peptide SEQ ID NO: 11.  
 DE  
 XX Cytochemical assay; quality control; immunohistochemical stain;  
 KW cancer; viral infection.  
 XX Synthetic.  
 OS  
 XX WO200062064-A2.  
 PN  
 XX 19-OCT-2000.  
 PD  
 XX 14-APR-2000; 2000WO-US10268.  
 PF  
 XX 14-APR-1999; 99US-0291351.  
 PR  
 XX (CYTO-) CYTOLOGIX CORP.  
 PA  
 XX Rogen SA, Radcliffe GE, Sompuram SR, Ramanathan H;  
 PI WPI; 2000-679514/66.  
 DR  
 XX

PT Quality control device for cytochemical assays measuring analytes in a  
 PT sample, comprises matrix with top surface having a quality control  
 PT reagent moiety and a bottom adhesive surface for adhering device to  
 PT platform  
 XX  
 PS Claim 20; Fig 6; 75pp; English.  
 XX  
 CC The present invention is related to a device useful for measuring the  
 CC quality of cytochemical assays such as immunohistochemical stains. These  
 CC assays can be used in cancer diagnosis, and in the diagnosis of viral  
 CC infections. The quality control device comprises a matrix with a top  
 CC surface having a quality control reagent moiety and the bottom surface  
 CC having an adhesive to enable the adherence of the device to a planar  
 CC test platform. The present sequence was used in assays to demonstrate the  
 CC device.  
 XX  
 SQ Sequence 18 AA;  
 Query Match 29.6%; Score 40; DB 21; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 60;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CNTSQDYTYCFN 12  
 DB 5 cdsnqtppycfn 16  
 ||::| ||||  
 RESULT 8  
 AAG98102  
 ID AAG98102 standard; Peptide; 14 AA.  
 XX  
 AC AAG98102;  
 XX 19-SEP-2001 (first entry)  
 DT  
 XX Human SNP associated peptide SEQ ID NO. 744.  
 DE  
 XX Human; single nucleotide polymorphism; SNP; angiopoietin;  
 KW 4-hydroxybutyrate; dehydrogenase; protein therapy;  
 KW adenosine triphosphate-dependent RNA helicase;  
 KW major histocompatibility complex Class I histocompatibility antigen; MHC;  
 KW phosphoglycerate kinase; immunosuppressive; antitumor; cytostatic;  
 KW antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;  
 KW antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine.  
 XX Homo sapiens.  
 OS  
 XX WO200148245-A2.  
 PN  
 XX 05-JUL-2001.  
 PD  
 XX 27-DEC-2000; 2000WO-US35346.  
 PF  
 XX 27-DEC-1999; 99US-0472688.  
 PR  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Shimkets RA, Leach M;  
 PI WPI; 2001-418297/44.  
 DR  
 XX Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,  
 PT adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate  
 PT kinase, useful for diagnosing and treating, e.g. cancer, autoimmune  
 PT diseases and infections  
 XX  
 PS Disclosure; Page 454; 484pp; English.  
 XX  
 CC The invention relates to nucleic acids (AAH79386-AAH80036) encoding  
 CC polymorphic variants of proteins (AAG98010-AAG98238) related to  
 CC angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate  
 CC (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)

CC Class I histocompatibility antigen and/or phosphoglycerate kinase. These  
 CC nucleic acid single nucleotide polymorphisms (SNPs) and the encoded  
 CC proteins have potential immunosuppressive, immunostimulatory,  
 CC antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,  
 CC antileukemic, neuroprotective and antimicrobial activity and may be  
 CC useful in gene/protein therapy, vaccines, modulation of the expression  
 CC and activity of proteins related to angiotensin, 4-hydroxybutyrate,  
 CC dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,  
 CC major histocompatibility complex (MHC) Class I histocompatibility antigen  
 CC and/or phosphoglycerate kinase. Disorders that may be prevented,  
 CC diagnosed and/or treated by the above methods include multifactorial  
 CC diseases with a genetic component, such as autoimmune diseases (e.g.  
 CC rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus  
 CC erythematosus and Grave's disease), inflammation, cancer (e.g. cancers  
 CC of the bladder, brain, breast, colon and kidney, leukemia), diseases of  
 CC the nervous system, an infection of pathogenic organisms. They may also  
 CC be used to alter phenotypic traits such as longevity, appearance,  
 CC strength, speed and endurance.

XX Sequence 14 AA;  
 SQ

Query Match 28.1%; Score 38; DB 22; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTSQYDT 8  
 |||||  
 Db 8 ntsqydt 14

RESULT 9  
 AAB48871  
 ID AAB48871 standard; peptide: 21 AA.

XX AC AAB48871;

XX DT 16-MAR-2001 (first entry)

XX DE Mutant human insulin (IA protein) trz\_06 A chain.

XX KW Human insulin; insulin activity protein; IA protein;  
 KW diabetes mellitus; type 1; type 2; stability;

XX KW mutant; mutin.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069901-A2.

XX PD 23-NOV-2000.

XX PF 19-MAY-2000; 2000WO-US13764.

XX PR 19-MAY-1999; 99US-0134930.

XX PA (XENC-) XENCOR INC.

XX PI Dahiyat BI;

XX WPI: 2001-025004/03.

PT Non-naturally occurring protein with insulin activity useful for  
 PT treating type 1 and type 2 diabetes, comprising amino acid  
 PT substitutions as compared to native human insulin and having enhanced  
 PT stability -

XX PS Claim 10; Fig 5A; 95pp; English.

XX CC The invention relates to novel non-naturally occurring mature human  
 CC insulin mutants, designated insulin activity (IA) proteins in the  
 CC specification (AAB48850-B48876), which have altered properties (e.g.,  
 CC enhanced oxidative, alkaline or thermal stability, prolonged shelf-life,

CC altered ability to oligomerise) compared to wild-type mature human  
 CC insulin (AAB48848, AAB48849) but which are still capable of binding to  
 CC an insulin receptor. The insulin mutants of the invention have less than  
 CC 98% sequence identity to wild-type human insulin, but have a 3D  
 CC structure which substantially corresponds to that of wild-type human  
 CC insulin. The invention also relates to recombinant nucleic acids encoding  
 CC the insulin mutants, expression vectors and host cells comprising mutant  
 CC insulin-encoding DNA, the recombinant preparation of an insulin mutant,  
 CC and a pharmaceutical composition comprising an insulin mutant of the  
 CC invention. The mature human insulin mutants are useful for treating  
 CC insulin-responsive conditions and disorders of carbohydrate metabolism  
 CC such as type 1 or type 2 diabetes. The nucleic acids encoding the  
 CC insulin mutants are used in gene therapy techniques for treating these  
 CC conditions. The insulin mutants are more stable compared to wild-type  
 CC mature human insulin, or have an improved ability to form insulin  
 CC hexamers. Sequences AAB48871-B48876 represent the A and B chains of  
 CC insulin mutants which have improved stability.

XX Sequence 21 AA;  
 SQ

Query Match 27.4%; Score 37; DB 22; Length 21;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CNTSQDYTYC 10  
 | : || : ||  
 Db 11. cslyqyenyc 20

RESULT 10

AAY85630

ID AAY85630 standard; protein: 17 AA.

XX AC AAY85630;

XX DT 07-FEB-2001 (first entry)

XX DE Snake venom derived protein fragment #3.

XX KW Subunit peptide production; snake venom; rattlesnake; thrombolytic;  
 KW von Willebrand's factor; blood platelet-inhibitory activity.

XX OS Crotalus horridus horridus.

XX PN WO200059926-A1.

XX PD 12-OCT-2000.

XX PF 31-MAR-2000; 2000WO-JP02127.

XX PR 02-APR-1999; 99JP-0096073.

XX PA (AJIN ) AJINOMOTO CO INC.

XX PI Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;

XX DR WPI: 2000-664985/64.

XX PT Producing physiologically-active subunit peptides originating in  
 PT polymer proteins by denaturation and specific separation, with lower  
 PT antigenicity but improved solubility and stability, e.g. blood  
 PT platelet-binding inhibitors -

XX PS Disclosure; Fig 6; 51pp; Japanese.

XX CC This invention relates to a method for the production of a subunit  
 CC peptide originating from a polymer protein with disulphide bonds within  
 CC and between subunits. The method comprises denaturing the protein or its  
 CC subunit using a protein denaturing agent in a solution, removing the  
 CC agent in the presence of a polyoxyalkyl polyether which reacts with a  
 CC thiol group and unwinds the subunit, and separating the polyoxyalkyl  
 CC polyether-bound subunit peptide. The method can be used for producing

CC physiologically-active subunit peptides for polymer proteins e.g. snake  
 CC venom-originated dimer peptide with blood platelet-inhibitory activity  
 CC on von Willebrand's factor. The peptides produced have platelet-binding  
 CC inhibitory, and thrombolytic activity. The present sequence represents a  
 CC fragment of a rattlesnake protein used in an example illustrating the  
 CC method of the invention.

XX SQ Sequence 17 AA;

Query Match 25.9%; Score 35; DB 21; Length 17;  
 Best Local Similarity 62.3%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 SOYDTYCF 11  
 I I I I I  
 Db 9 ssydrycy 16

RESULT 11  
 AAY98406  
 ID AAY98406 standard; Peptide; 13 AA.

XX AC AAY98406;

XX DT 31-JUL-2000 (first entry)

XX DE Alpha D peptide derived from FGFR-3 SEQ ID NO:121.

XX KW Alpha D peptide; Alpha D region; protein kinase; modulation; activity;  
 KW cytosolic; anti-diabetic; anorectic; antiinflammatory; dermatological;  
 KW immunosuppressive; immunomodulator; osteopathic; cardiant; vasotropic;  
 KW antiarteriosclerotic; protein kinase modulator; cancer; proliferation;  
 KW restenosis; atherosclerosis; skin disorder; diabetes; obesity;  
 KW central nervous system disorder; inflammatory disorder; osteoporosis;  
 KW autoimmune disease; immune disorder; cardiovascular disease.

XX OS Homo sapiens.

XX PN WO200018895-A1.

XX PD 06-APR-2000.

XX PF 24-SEP-1999; 99WO-US221106.

XX PR 25-SEP-1998; 98US-0161094.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX PT Ben-Sasson SA;

XX DR WPI; 2000-328722/28.

XX PT Peptide derivatives of protein kinase alpha D regions which selectively  
 modulate the activity of protein kinases -

XX PS Claim 23; Fig 3; 148pp; English.

XX CC The present invention describes a peptide derivative (A) of the protein  
 CC kinase alpha D region comprising 5-30 amino acids, which modulates  
 CC the activity of the protein kinase. AAY98286 to AAY98455 represent  
 CC peptides derived from protein kinase alpha D regions, which are used in  
 CC the exemplification of the present invention. The peptides have  
 CC cytosolic, anti-diabetic, anorectic, antiinflammatory, dermatological,  
 CC cardiant, immunosuppressive, immunomodulator, osteopathic, vasotropic,  
 CC and antiarteriosclerotic activities, and are protein kinase modulators.  
 CC The peptides can be used as test peptides to identify protein kinase  
 CC modulators. They can also be used to modulate the activity of a protein  
 CC kinase in a subject, and in a method of detecting a ligand that binds  
 CC to the alpha D region of a protein kinase. They may be used to  
 CC produce antibodies that bind to the alpha D region of a protein kinase.  
 CC The peptides are useful in the treatment of diseases caused by over-

CC or under-activity of a protein kinase, e.g. cancer, diseases caused by  
 CC proliferation of smooth muscle (e.g. restenosis and atherosclerosis),  
 CC skin disorders, diabetes, obesity, diseases of the central nervous  
 CC system, inflammatory disorders, autoimmune diseases and other immune  
 CC disorders, osteoporosis and cardiovascular diseases.

XX SQ Sequence 13 AA;

Query Match 25.2%; Score 34; DB 21; Length 13;  
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 FNASAPPEE 19  
 I : I I I I I  
 Db 3 fdtskppee 11

RESULT 12  
 AAJ03726  
 ID AAJ03726 standard; Peptide; 15 AA.

XX AC AAJ03726;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #3717.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
 KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US19774.

XX PR 19-JUL-1999; 99US-0357737.

XX PA (EPIM-) EPIMUNE INC.

PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

PI Baker DM, Cellis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccines against hepatitis C virus -

XX PS Disclosure; Page 182; 214pp; English.

XX CC The present invention describes a composition comprising a prepared  
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.  
 CC These are derived from HCV HLA-binding motifs. They are useful in  
 CC vaccines for the prevention and treatment of HCV infection in humans. The  
 CC present sequence is an epitope used in the disclosure of the invention.

XX SQ Sequence 15 AA;

Query Match 25.2%; Score 34; DB 22; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 YDTYCFNAS 14  
 I I I I I I I I I I  
 Db 6 ydtrcfst 14

RESULT 13  
 AAJ03739  
 ID AAJ03739 standard; Peptide; 15 AA.

XX AAJ03739;  
 XX  
 DT 02-JUL-2001 (first entry)  
 XX  
 DE Hepatitis C virus epitope #3730.  
 XX  
 KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
 KW antiviral.  
 XX  
 OS Hepatitis C virus.  
 XX  
 XX WO200121189-A1.  
 XX  
 PD 29-MAR-2001.  
 XX  
 XX 19-JUL-2000; 2000WO-US19774.  
 PF  
 XX 19-JUL-1999; 99US-0357737.  
 PR  
 XX (EPIM-) EPIIMUNE INC.  
 PA  
 XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;  
 XX  
 XX WPI; 2001-308046/32.  
 DR  
 XX A new composition useful as a vaccines against hepatitis C virus -  
 PT  
 XX Disclosure; Page 183; 214pp; English.  
 PS  
 XX The present invention describes a composition comprising a prepared  
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.  
 CC These are derived from HCV HLA-binding motifs. They are useful in  
 CC vaccines for the prevention and treatment of HCV infection in humans. The  
 CC present sequence is an epitope used in the disclosure of the invention.  
 XX  
 XX Sequence 15 AA;  
 SQ

Query Match 25.2%; Score 34; DB 22; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 YDTYCFNAS 14  
 III III:  
 DB 6 ydtcrfdst 14

RESULT 14  
 AA42772  
 ID AA42772 standard; Peptide; 21 AA.  
 XX  
 AC AA42772;  
 XX  
 DT 27-APR-1998 (first entry)  
 XX  
 DE Peptide derived from Spl7.  
 XX  
 XX Spl7; sperm autoantigen; antigen; antigenic epitope; fertility;  
 KW immunocontraception; immunocontraceptive vaccine; autoimmune fertility.  
 KW  
 XX Synthetic.  
 OS  
 XX WO9739020-A2.  
 PN  
 XX 23-OCT-1997.  
 PD  
 XX 11-APR-1997; 97WO-US06489.  
 PF  
 XX 15-APR-1996; 96US-0632535.  
 PR  
 XX (UYN-) UNIV NORTH CAROLINA.  
 PA

XX Lea I, O'Rand MG, Widgren EE;  
 PI  
 XX WPI; 1997-526394/48.  
 DR  
 XX Sperm protein Spl7 antigenic peptide(s) - used as  
 PT immunocontraceptive to reduce fertility  
 XX  
 XX Claim 1; Page 83; 160pp; English.  
 PS  
 XX The present peptide is derived from a Spl7 protein. Spl7 peptides contain  
 CC antigenic epitopes in mice, rabbits, non-human primates and humans. They  
 CC induce antibodies which recognise sperm, and inhibit fertilisation. The  
 CC peptides can be used in immunocontraceptive methods as a  
 CC immunocontraceptive vaccine to reduce the fertility of animals,  
 CC particularly in females. The presence of the antigenic peptides can be  
 CC detected to diagnose autoimmune fertility in both male and female  
 CC subjects.  
 XX  
 XX Sequence 21 AA;  
 SQ

Query Match 25.2%; Score 34; DB 18; Length 21;  
 Best Local Similarity 35.7%; Pred. No. 4.4e+02;  
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 6 YDTYCFNASAPPEE 19  
 I: : I III:  
 DB 8 ynnhafaeeqpepek 21

RESULT 15  
 AAB82056  
 ID AAB82056 standard; peptide; 20 AA.  
 XX  
 AC AAB82056;  
 XX  
 DT 22-JUN-2001 (first entry)  
 DT  
 XX HCV antigen, NS3 1585-1604.  
 DE  
 XX Antigen; immunostimulant; vaccine; pharmaceutical composition; antiviral;  
 KW viral infection.  
 KW  
 XX Hepatitis C virus.  
 OS  
 XX WO200124822-A2.  
 PN  
 XX 12-APR-2001.  
 PD  
 XX 02-OCT-2000; 2000WO-EP09657.  
 PF  
 XX 01-OCT-1999; 99AT-0001680.  
 PR  
 XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.  
 PA  
 XX Fleitmann J, Mattner F, Buschle M, Melling J;  
 PI  
 XX WPI; 2001-290577/30.  
 DR  
 XX New pharmaceutical composition comprising an antigen, an  
 PT immunostimulating substance and a polycationic polymer, useful in  
 PT manufacturing vaccines -  
 XX  
 XX Claim 12; Page 16; 20pp; English.  
 PS  
 XX The present invention relates to a pharmaceutical composition comprising  
 CC (a) an antigen; (b) an immunostimulating substance consisting of  
 CC neuroactive compounds, hormones, compounds having growth hormone activity  
 CC or their mixtures; and (c) a polycationic polymer. The present sequence  
 CC is an antigenic peptide derived from Hepatitis C virus, which was used in  
 CC the present invention. The composition is useful in manufacturing  
 CC vaccines.

XX  
SQ Sequence 20 AA;

Query Match 24.4%; Score 33; DB 22; Length 20;  
Best Local Similarity 60.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 TYCFNASAPP 17  
| | | | |  
Db 8 tvcaragapp 17

Search completed: February 21, 2002, 16:30:34  
Job time: 175. sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:28:59 ; Search time 12.63 Seconds  
(without alignments)  
40.980 Million cell updates/sec

Title: US-08-753-851-2  
Perfect score: 135  
Sequence: 1 CNTSQYDTYCFNASAPPEEDCTS 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 119836

```
Minimum DB seq length: 0
Maximum DB seq length: 23
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```

Database :
Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	135	100.0	23	2	US-08-143-311B-2	Sequence 2, Appl
2	45	33.3	11	2	US-08-143-311B-14	Sequence 14, Appl
3	33.5	24.8	22	1	US-08-179-481-108	Sequence 108, App
4	32	23.7	14	1	US-08-433-318A-104	Sequence 104, App
5	32	23.7	14	1	US-08-433-318A-105	Sequence 105, App
6	32	23.7	14	1	US-08-433-318A-167	Sequence 167, App
7	32	23.7	14	1	US-08-433-318A-168	Sequence 168, App
8	32	23.7	14	2	US-08-922-048-104	Sequence 104, App
9	32	23.7	14	2	US-08-922-048-105	Sequence 105, App
10	32	23.7	14	2	US-08-922-048-167	Sequence 167, App
11	32	23.7	14	2	US-08-922-048-168	Sequence 168, App
12	32	23.7	14	5	PCT-US96-06270-104	Sequence 104, App
13	32	23.7	14	5	PCT-US96-06270-105	Sequence 105, App
14	32	23.7	14	5	PCT-US96-06270-167	Sequence 167, App
15	32	23.7	14	5	PCT-US96-06270-168	Sequence 168, App
16	32	23.7	15	1	US-08-433-318A-88	Sequence 88, App
17	32	23.7	15	1	US-08-433-318A-89	Sequence 89, Appl
18	32	23.7	15	1	US-08-433-318A-151	Sequence 151, App
19	32	23.7	15	1	US-08-433-318A-152	Sequence 152, App
20	32	23.7	15	2	US-08-922-048-88	Sequence 88, Appl
21	32	23.7	15	2	US-08-922-048-89	Sequence 89, Appl
22	32	23.7	15	2	US-08-922-048-151	Sequence 151, App
23	32	23.7	15	2	US-08-922-048-152	Sequence 152, App
24	32	23.7	15	5	PCT-US96-06270-88	Sequence 88, Appl
25	32	23.7	15	5	PCT-US96-06270-89	Sequence 89, Appl
26	32	23.7	15	5	PCT-US96-06270-151	Sequence 151, App
27	32	23.7	15	5	PCT-US96-06270-152	Sequence 152, App

28	31.5	23.3	11	4	US-08-866-545-16	Sequence 16,	Appl
29	31.5	23.3	18	1	US-08-179-481-99	Sequence 99,	Appl
30	31.5	23.3	19	1	US-08-520-977A-8	Sequence 8,	Appl
31	31	23.0	11	1	US-07-958-903A-9	Sequence 9,	Appl
32	31	23.0	11	1	US-08-462-018-9	Sequence 9,	Appl
33	31	23.0	11	1	US-08-823-245-9	Sequence 9,	Appl
34	31	23.0	11	5	PCT-US93-09443A-59	Sequence 59,	Appl
35	31	23.0	14	4	US-08-853-910-8	Sequence 8,	Appl
36	30.5	22.6	16	1	US-07-958-903A-6	Sequence 6,	Appl
37	30.5	22.6	16	1	US-08-462-018-6	Sequence 6,	Appl
38	30.5	22.6	16	1	US-08-823-245-6	Sequence 6,	Appl
39	30.5	22.6	16	5	PCT-US92-09443A-17	Sequence 17,	Appl
40	30	22.2	10	1	US-08-433-318A-136	Sequence 136,	App
41	30	22.2	10	1	US-08-433-318A-137	Sequence 137,	App
42	30	22.2	10	2	US-08-922-048-136	Sequence 136,	App
43	30	22.2	10	2	US-08-922-048-137	Sequence 137,	App
44	30	22.2	10	5	PCT-US96-06270-136	Sequence 136,	App
45	30	22.2	10	5	PCT-US96-06270-137	Sequence 137,	App

## ALIGNMENTS

```

1
RESULT
US-08-143-311B-2
Sequence 2, Application US/08143311B
Patent No. 5863540
GENERAL INFORMATION:
APPLICANT: HAYNES, BARTON F.
APPLICANT: HALE, LAURA P.
APPLICANT: PATTON, KAREN L.
APPLICANT: TELEN, MARILYN J.
APPLICANT: LIAO, HUA-XIN
TITLE OF INVENTION: AN ADHESION MOLECULE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/143,311B
FILING DATE: 29-OCT-1993
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,339
FILING DATE: 30-OCT-1992
CLASSIFICATION: 436
APPLICATION NUMBER: 07/669,730
FILING DATE: 15-MAR-1991
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-143-311B-2

```

### RESULT 3

/ GENERAL INFORMATION:  
 / APPLICANT: Greene, Mark I.  
 / TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
 / TITLE OF INVENTION: METHODS OF USING THE SAME  
 / NUMBER OF SEQUENCES: 184  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSER: Woodcock Washburn Kurtz Mackiewicz &  
 / ADDRESSER: No. 5663144ris  
 / STREET: One Liberty Place - 46th Floor  
 / CITY: Philadelphia  
 / STATE: PA  
 / COUNTRY: USA  
 / ZIP: 19103  
 / COMPUTER READABLE FORM:

### RESULT 3

;; MEDIUM TYPE: DISKETTE, 3.5 INCH  
;; COMPUTER: IBM PC Compatible  
;; OPERATING SYSTEM: Windows  
;; SOFTWARE: WORDPERFECT 6  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/433,318A  
;; FILING DATE: 03-MAY-1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mark DeLuca  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: UPN-2106  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 104:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
US-08-433-318A-104

Query Match 23.7%; Score 32; DB 1; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12

RESULT 5  
US-08-433-318A-105  
;; Sequence 105, Application US/08433318A  
;; Patent No. 5663144  
;; GENERAL INFORMATION:  
;; APPLICANT: Greene, Mark I.  
;; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
;; TITLE OF INVENTION: METHODS OF USING THE SAME  
;; NUMBER OF SEQUENCES: 184  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
;; STREET: One Liberty Place - 46th Floor  
;; CITY: Philadelphia  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19103  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: DISKETTE, 3.5 INCH  
;; COMPUTER: IBM PC Compatible  
;; OPERATING SYSTEM: Windows  
;; SOFTWARE: WORDPERFECT 6  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/433,318A  
;; FILING DATE: 03-MAY-1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mark DeLuca  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: UPN-2106  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 105:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
US-08-433-318A-105

Query Match 23.7%; Score 32; DB 1; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12

RESULT 6  
US-08-433-318A-167  
;; Sequence 167, Application US/08433318A  
;; Patent No. 5663144  
;; GENERAL INFORMATION:  
;; APPLICANT: Greene, Mark I.  
;; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
;; TITLE OF INVENTION: METHODS OF USING THE SAME  
;; NUMBER OF SEQUENCES: 184  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
;; STREET: One Liberty Place - 46th Floor  
;; CITY: Philadelphia  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19103  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: DISKETTE, 3.5 INCH  
;; COMPUTER: IBM PC Compatible  
;; OPERATING SYSTEM: Windows  
;; SOFTWARE: WORDPERFECT 6  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/433,318A  
;; FILING DATE: 03-MAY-1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mark DeLuca  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: UPN-2106  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 167:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
US-08-433-318A-167

Query Match 23.7%; Score 32; DB 1; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12

RESULT 7  
US-08-433-318A-168



; Sequence 168, Application US/08433318A  
; Patent No. 5663144  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5663144ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WORDPERFECT 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/433,318A  
; FILING DATE: 03-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark Deluca  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: UPN-2106  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 168:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-433-318A-168

Query Match 23.7%; Score 32; DB 1; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 1;

QY 5 QYDYCFNASAPP 17  
|:| |:  
Db 4 QWDWYCY---PP 12

RESULT 8  
US-08-922-048-104  
; Sequence 104, Application US/08922048  
; Patent No. 5919764  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5919764ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows

; SOFTWARE: WORDPERFECT 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/922,048  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/433,318  
; FILING DATE: 03-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark Deluca  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: UPN-2106  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 104:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-922-048-104

Query Match 23.7%; Score 32; DB 2; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 QYDYCFNASAPP 17  
|:| |:  
Db 4 QWDWYCY---PP 12

RESULT 9  
US-08-922-048-105  
; Sequence 105, Application US/08922048  
; Patent No. 5919764  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5919764ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WORDPERFECT 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/922,048  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/433,318  
; FILING DATE: 03-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark Deluca  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: UPN-2106  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14  
; TYPE: amino acid

TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-922-048-105

Query Match 23.7%; Score 32; DB 2; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 5 QYDTYCFNASAPP 17  
Db 4 QMDWYCY----PP 12

## RESULT 10

US-08-922-048-167  
Sequence 167, Application US/08922048  
Patent No. 5919764

GENERAL INFORMATION:  
APPLICANT: Greene, Mark I.

TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 5919764ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

## COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WORDPERFECT 6  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/922.048  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/433,318  
FILING DATE: 03-MAY-1995

## ATTORNEY/AGENT INFORMATION:

NAME: Mark Deluca  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2106  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 167:

## SEQUENCE CHARACTERISTICS:

LENGTH: 14

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-922-048-167

Query Match 23.7%; Score 32; DB 2; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 5 QYDTYCFNASAPP 17  
Db 4 QMDWYCY----PP 12

## RESULT 11

US-08-922-048-168

Sequence 168, Application US/08922048  
Patent No. 5919764

GENERAL INFORMATION:

APPLICANT: Greene, Mark I.  
TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 5919764ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

## COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WORDPERFECT 6  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/922,048  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/433,318  
FILING DATE: 03-MAY-1995

## ATTORNEY/AGENT INFORMATION:

NAME: Mark Deluca  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 168:

## SEQUENCE CHARACTERISTICS:

LENGTH: 14

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-922-048-168

Query Match 23.7%; Score 32; DB 2; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 5 QYDTYCFNASAPP 17  
Db 4 QMDWYCY----PP 12

## RESULT 12

PCT-US96-06270-104

Sequence 104, Application PC/TUS9606270  
GENERAL INFORMATION:

APPLICANT: Greene, Mark I.

APPLICANT: Zhang, Xin

TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 184

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

## COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/06270  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/433,318  
FILING DATE: 03-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark DeLuca  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2748  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US96-06270-104

Query Match 23.7%; Score 32; DB 5; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12

## RESULT 13

PCT-US96-06270-105  
SEQUENCE 105, Application PC/TUS9606270  
GENERAL INFORMATION:  
APPLICANT: Greene, Mark I.  
TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06270  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/433,318  
FILING DATE: 03-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark DeLuca  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2748  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US96-06270-105

Query Match 23.7%; Score 32; DB 5; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12  
RESULT 14  
PCT-US96-06270-167  
SEQUENCE 167, Application PC/TUS9606270  
GENERAL INFORMATION:  
APPLICANT: Greene, Mark I.  
TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06270  
FILING DATE:

QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/433,318  
FILING DATE: 03-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark DeLuca  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2748  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US96-06270-105

Query Match 23.7%; Score 32; DB 5; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWYCY----PP 12  
RESULT 15  
PCT-US96-06270-168  
SEQUENCE 168, Application PC/TUS9606270  
GENERAL INFORMATION:  
APPLICANT: Greene, Mark I.  
TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris

STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06270  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/433,318  
FILING DATE: 03-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark DeLuca  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2748  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 168:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US96-06270-168

Query Match 23.7%; Score 32; DB 5; Length 14;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
QY 5 QYDTYCFNASAPP 17  
Db 4 QWDWICY----PP 12

Search completed: February 21, 2002, 16:30:53  
Job time: 114 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:33:15 ; Search time 12.69 seconds  
(without alignments)  
192.087 Million cell updates/sec

Title: US-08-753-851-3  
Perfect score: 179  
Sequence: 1 CRDGRVQKGEYRTNPEDIYPSNPTDDVSS 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 6754

Minimum DB seq length: 0  
Maximum DB seq length: 32

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	22.3	29	2	I49732 NADH dehydrogenase
2	35	19.6	25	2	T01689 ATPase-beta chain
3	35	19.6	30	2	PC2307 X-pro aminopeptidase
4	34	19.0	27	2	S11176 aspartate transcarbamoylase
5	34	19.0	29	2	S35924 T-cell receptor gamma
6	33	18.4	19	2	A41299 T-cell receptor alpha
7	33	18.4	28	2	A03356 omega-gliadin - epsilon
8	31.5	17.6	20	2	S44465 pyrrolic acid - pyruvate
9	31.5	17.6	30	2	A84412 hypothetical protein
10	31	17.3	17	2	C37396 pollen allergen Fe
11	30	16.8	20	2	G61491 seed protein ws-25
12	30	16.8	20	2	D37396 pollen allergen Fe
13	30	16.8	22	2	I50533 calmodulin - epsilon
14	30	16.8	26	2	F37396 pollen allergen Po
15	30	16.8	26	2	E37396 pollen allergen Ag
16	30	16.8	26	2	G37396 pollen allergen An
17	30	16.8	29	2	B56817 photosystem I chain
18	30	16.8	29	2	B60698 trichocyst protein
19	30	16.8	29	2	A11683 hyaluronate receptor
20	29.5	16.5	19	2	A61377 endometrial secret
21	29	16.2	24	2	A39509 mannose-specific lectin
22	29	16.2	25	2	S28837 RNA-binding protein
23	29	16.2	26	2	D53507 annexin V - rat (f)
24	29	16.2	28	2	PS0106 2-phosphogluconate
25	29	16.2	31	2	T01701 hypothetical protein
26	29	16.2	32	4	B34962 hypothetical protein
27	28	15.6	18	2	C24780 protein-tyrosine kinase
28	28	15.6	24	2	I54329 gene NF1 protein - cDNA
29	28	15.6	25	2	C25629 cytochrome-c oxidase

30 27.5 15.4 20 2 S32387 ribosomal protein  
31 27 15.1 15 2 S27248 pseudogerm - whe  
32 27 15.1 15 2 C44101 calmodulin, vasoac  
33 27 15.1 20 2 A54519 tubulin alpha chain  
34 27 15.1 22 2 S13899 glycoprotein, 11.7  
35 27 15.1 22 2 S52357 hypothetical prote  
36 27 15.1 25 2 P00369 L protein - rabies  
37 27 15.1 26 2 S51055 ribosomal protein  
38 27 15.1 32 2 C49480 major immunophilin  
39 26 14.5 14 1 LFECW trp operon leader  
40 26 14.5 15 2 B85761 chemottractant pr  
41 26 14.5 15 2 A36279 calsequestrin, car  
42 26 14.5 18 2 A39040 protein phosphatas  
43 26 14.5 18 2 A32917 major outer membra  
44 26 14.5 20 2 S28434 cytochrome-c oxida  
45 26 14.5 20 2 S77983

#### ALIGNMENTS

##### RESULT 1

I49732  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) 14K chain - potato mitochondrion (fragme  
N:Alternate names: complex 1 dehydrogenase 14K chain; NADH-ubiquinone oxidoreductase  
C:Species: mitochondrion Solanum tuberosum (potato)  
C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 07-May-1999  
C:Accession: I49732  
R:Herz, U.; Schroeder, W.; Liddell, A.; Leaver, C.J.; Brennicke, A.; Grohmann, L.  
J. Biol. Chem. 269, 2263-2269, 1994  
A:Title: Purification of the NADH:ubiquinone oxidoreductase (complex I) of the respir  
A:Reference number: A49732; MUID:94124587  
A:Accession: I49732  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-29 <HER>  
C:Genetics:  
A:Genome: mitochondrion  
C:Keywords: electron transfer; hydrogen ion transport; mitochondrial inner membrane;

Query Match 22.3%; Score 40; DB 2; Length 29;  
Best Local Similarity 46.7%; Pred. No. 45;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 17 PEDIYPSNPTDDVVS 31  
||| :||| ||:  
Db 13 PDDFDPNNPYKKDVA 27

##### RESULT 2

T01689  
ATPase-beta chain - maize chloroplast (fragment)  
C:Species: chloroplast Zea mays (maize)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 29-Oct-1999  
C:Accession: T01689  
R:Bradley, D.; Gatenby, A.A.  
EMBO J. 4, 3641-3648, 1985  
A:Title: Mutational analysis of the maize chloroplast ATPase-beta subunit gene promot  
system.  
A:Reference number: Z14396; MUID:86136006  
A:Accession: T01689  
A:Status: preliminary; translated from GB/EMBL/DBDJB  
A:Molecule type: DNA  
A:Residues: 1-25 <BRA>  
A:Cross-references: EMBL:X03396; NID:g12404; PIDN:CAA27131.1; PID:g12405  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 19.6%; Score 35; DB 2; Length 25;  
Best Local Similarity 38.9%; Pred. No. 1.9e+02;

Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 14 RTNPDIYPSNPPTDDVYS 131  
||||| :| :|  
Db 2 RTNPITSRPGDPLESTCS 19

## RESULT 3

PC2307  
X-Pro aminopeptidase (EC 3.4.11.9) L13K - guinea pig (fragment)  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 25-Feb-1995 #sequence\_revision 26-May-1995 #text\_change 17-Nov-2000  
C:Accession: PC2307; PC2310  
R:Denslow, N.D.; Ryan, J.W.; Nguyen, H.P.  
Biochem. Biophys. Res. Commun. 205, 1790-1795, 1994  
A:Title: Guinea pig membrane-bound aminopeptidase P is a member of the proline peptidase  
A:Reference number: PC2307; MUID:95110325  
A:Accession: PC2307  
A:Molecule type: protein  
A:Residues: 1-30 <DEN>  
A:Experimental source: lung  
A:Accession: PC2310  
A:Molecule type: protein  
A:Residues: 1-26 <DE2>  
A:Experimental source: kidney  
C:Comment: This enzyme is a membrane bound protein.  
C:Comment: This enzyme hydrolyzes imido bonds.  
C:Comment: This enzyme acts as an aminocarpoxylproline hydrolase and is a member of the proline  
C:Superfamily: pig x-pro aminopeptidase  
C:Keywords: alpha-aminoacylpeptide hydrolase; membrane protein

Query Match 19.6%; Score 35; DB 2; Length 30;  
Best Local Similarity 40.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 7 YVQKGEYRTNPDIY 21  
| | | | :| :|  
Db 10 YYQDGEFGIRIEDVF 24

## RESULT 4

S51176  
aspartate transcarbamoylase - hamster  
C:Species: Cricetinae gen. sp. (hamster)  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
C:Accession: S51176  
R:Hemmens, B.; Carrey, E.A.  
Eur. J. Biochem. 225, 845-853, 1994  
A:Title: Proteolytic cleavage of the multienzyme polypeptide CAD to release the mammalian  
A:Reference number: S51176; MUID:95045561  
A:Accession: S51176  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-27 <HEM>

Query Match 19.0%; Score 34; DB 2; Length 27;  
Best Local Similarity 44.4%; Pred. No. 2.8e+02;  
Matches 8; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 12 EYKTNPET---DIYSPNP 25  
| | | | :| :|  
Db 3 EITTPETGTPDPYPASP 20

## RESULT 5

S35924  
T-cell receptor gamma chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 23-Mar-2001  
C:Accession: S35924  
R:Mathioudakis, G.; Platsoucas, C.D.

submitted to the EMBL Data Library, May 1993  
A:Description: An alternative splicing between VgammaJgamma2.3 and Cgamma2 gene segments  
A:Reference number: S32764  
A:Accession: S35924  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-29 <MAT>  
A:Cross-references: EMBL:Z22684  
C:Keywords: T-cell receptor

Query Match 19.0%; Score 34; DB 2; Length 29;  
Best Local Similarity 53.8%; Pred. No. 3e+02;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GTRYVQKGEYRTN 16  
| | | | :| :|  
Db 12 GTTLVVTGKYRKN 24

## RESULT 6

A41299  
T-cell receptor alpha chain precursor V region (17.2) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 23-Jul-1999  
C:Accession: A41299  
R:Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panay, G.  
Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991  
A:Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumatoid  
A:Reference number: A41299; MUID:92020887  
A:Accession: A41299  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-19 <DEM>  
A:Cross-references: GB:S57433; NID:g236318; PIDN:AAB19956.1; PID:g236319  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor

Query Match 18.4%; Score 33; DB 2; Length 19;  
Best Local Similarity 58.3%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 19 DIVPSNPPTDDV 30  
| | | | :| :|  
Db 4 DIVPSQPGDSAV 15

## RESULT 7

A03356  
omega-gliadin - einkorn wheat (fragment)  
C:Species: Triticum monococcum (einkorn wheat)  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 18-Jun-1993  
C:Accession: A03356  
R:Shewry, P.R.; Autran, J.C.; Nimmo, C.C.; Lew, E.J.L.; Kasarda, D.D.  
Nature 286, 520-522, 1980  
A:Reference number: A93228  
A:Accession: A03356  
A:Molecule type: protein  
A:Residues: 1-28 <SHE>  
C:Superfamily: gliadin

Query Match 18.4%; Score 33; DB 2; Length 28;  
Best Local Similarity 40.0%; Pred. No. 4e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 16 NPEDIYPSNP 25  
| | | | :| :|  
Db 13 SPQQLYPOQP 22

## RESULT 8

S44465

pyrrhocoricin - Pyrrhocoris apterus  
 C:Species: Pyrrhocoris apterus  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-May-1999  
 C:Accession: S44465  
 R:Cociancich, S.; Dupont, A.; Hegy, G.; Lanot, R.; Holder, F.; Hetru, C.; Hoffmann, J.A.  
 Biochem. J. 300, 567-575, 1994  
 A:Title: Novel inducible antibacterial peptides from a hemipteran insect, the sap-sucking  
 A:Reference number: S44463; MUID:94271176  
 A:Accession: S44465  
 A:Molecule type: protein  
 A:Residues: 1-20 <COC>  
 C:Function:  
 A:Description: antibacterial protein  
 A:Note: active against Gram-negative bacteria  
 C:Keywords: antibacterial; hemolymph; immune response

Query Match 17.6%; Score 31.5; DB 2; Length 20;  
 Best Local Similarity 45.0%; Pred. No. 4.4e+02;  
 Matches 9; Conservative 0; Mismatches 8; Indels 3; Gaps 1;

QY 8 VQKGEY---RTNPEDIYPSN 24  
 | | | | | | | | | |  
 Db 1 VDKGSYLPRTPRPYINRN 20

RESULT 9

A84412  
 hypothetical protein Vng2626h [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: A84412  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
 ; Leitthausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84412  
 A:Accession: A84412  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-30 <STO>  
 A:Cross-references: GB:AE004437; NID:gi10582006; PIDN:AAG20661.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG2626H

Query Match 17.6%; Score 31.5; DB 2; Length 30;  
 Best Local Similarity 33.3%; Pred. No. 6.9e+02;  
 Matches 8; Conservative 2; Mismatches 9; Indels 5; Gaps 1;

QY 10 KGEYTNP-----EDIYPSNPTDD 28  
 : | | | | | | | | | |  
 Db 7 RGLLYTPTDSDQSDTTAQPADD 30

RESULT 10

C37396  
 pollen allergen Fes e I type A - reed fescue (fragment)  
 C:Species: Festuca elatior (reed fescue)  
 C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 25-Oct-1996  
 C:Accession: C37396  
 R:Esch, R.E.; Klapper, D.G.  
 Mol. Immunol. 26, 557-561, 1989  
 A:Title: Isolation and characterization of a major cross-reactive grass group I allergen  
 A:Reference number: A37396; MUID:89364850  
 A:Accession: C37396  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-17 <ESC>  
 C:Keywords: pollen

Query Match 17.3%; Score 31; DB 2; Length 17;  
 Best Local Similarity 31.2%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 7 YVOKGEYRTNPEDIYP 22  
 | | | | | | | | | |  
 Db 1 YTTGGTKSEADVIP 16

RESULT 11

G61491  
 seed protein ws-25 - winged bean (fragment)  
 C:Species: Psophocarpus tetragonolobus (winged bean)  
 C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 17-Mar-2000  
 C:Accession: G61491  
 R:Hirano, H.

J. Protein Chem. 8, 115-130, 1989  
 A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-  
 A:Reference number: A61491; MUID:89351606  
 A:Accession: G61491  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <HIR>  
 C:Superfamily: plant Kunitz-type proteinase inhibitor  
 C:Keywords: glycoprotein; seed

Query Match 16.8%; Score 30; DB 2; Length 20;  
 Best Local Similarity 45.5%; Pred. No. 7e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGTRYVOKGEY 13  
 | | | | | | | | | |  
 Db 9 DGNKLVNRGY 19

RESULT 12

D37396  
 pollen allergen Fes e I type B - reed fescue (fragment)  
 C:Species: Festuca elatior (reed fescue)  
 C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 09-Jun-2000  
 C:Accession: D37396  
 R:Esch, R.E.; Klapper, D.G.  
 Mol. Immunol. 26, 557-561, 1989  
 A:Title: Isolation and characterization of a major cross-reactive grass group I aller  
 A:Reference number: A37396; MUID:89364850  
 A:Accession: D37396  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <ESC>  
 C:Superfamily: expansin  
 C:Keywords: pollen

Query Match 16.8%; Score 30; DB 2; Length 20;  
 Best Local Similarity 31.2%; Pred. No. 7e+02;  
 Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 7 YVOKGEYRTNPEDIYP 22  
 | | | | | | | | | |  
 Db 1 YTTGGTKSEADVIP 16

RESULT 13

I50533  
 calmodulin - electric eel (fragment)  
 C:Species: Electrophorus electricus (electric eel)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 29-Sep-1999  
 C:Accession: I50533  
 R:Munjaal, R.P.; Dedman, J.R.; Means, A.R.  
 Ann. N. Y. Acad. Sci. 356, 110-118, 1980

A:Title: Isolation of the structural gene for calmodulin.

A:Reference number: I50533; MUID:81205530

A:Accession: I50533

A>Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 1-22 <MUN>

A:Cross-references: GB:M25058; NID:g213127; PIDN:AAA62797.1; PID:g213128

C:Superfamily: calmodulin; calmodulin repeat homology

C:Keywords: EF hand

Search completed: February 21, 2002, 16:35:06  
Job time: 111 sec

Query Match 16.8%; Score 30; DB 2; Length 22;  
Best Local Similarity 38.5%; Pred. NO. 7.8e+02;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 RDGTRYVQGEYR 14

Db 3 KONGYISAAELR 15

#### RESULT 14

E37396

pollen allergen Poa p I - Kentucky bluegrass (fragment)

C:Species: Poa pratensis (Kentucky bluegrass)

C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 09-Jun-2000

C:Accession: E37396

R:Esch, R.E.; Klapper, D.G.

Mol. Immunol. 26, 557-561, 1989

A:Title: Isolation and characterization of a major cross-reactive grass group I allergen

A:Reference number: A37396; MUID:89364850

A:Accession: E37396

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-26 <ESC>

C:Superfamily: expansin

C:Keywords: pollen

Query Match 16.8%; Score 30; DB 2; Length 26;  
Best Local Similarity 31.2%; Pred. NO. 9.4e+02;  
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 YVQKGEYRTNPEDIYP 22

Db 1 YTEGGTKAEADVIP 16

#### RESULT 15

E37396

pollen allergen Agr a I - bent grass (fragment)

C:Species: Agrostis alba (bent grass)

C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 09-Jun-2000

C:Accession: E37396

R:Esch, R.E.; Klapper, D.G.

Mol. Immunol. 26, 557-561, 1989

A:Title: Isolation and characterization of a major cross-reactive grass group I allergen

A:Reference number: A37396; MUID:89364850

A:Accession: E37396

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-26 <ESC>

C:Superfamily: expansin

C:Keywords: pollen

Query Match 16.8%; Score 30; DB 2; Length 26;  
Best Local Similarity 31.2%; Pred. NO. 9.4e+02;  
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 YVQKGEYRTNPEDIYP 22

Db 1 YTEGGTKAEADVIP 16





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:34:50 ; Search time 10.07 Seconds  
(without alignments)  
116.512 Million cell updates/sec

Title: US-08-753-851-3  
Perfect score: 179  
Sequence: 1 CRDGRYVQGEYRTNPDIPSNPTDDVSS 32

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1983

Minimum DB seq length: 0

Maximum DB seq length: 32

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	40	22.3	29	1 NU01_SOLTU
2	34	19.0	30	1 CY35_DESAC
3	33	18.4	24	1 FEDG_AMEYE
4	33	18.4	28	1 GDO_TRIMO
5	31.5	17.6	20	1 PYRR_PVRAP
6	30	16.8	29	1 PSAP_SYNP6
7	29.5	16.5	16	1 MK2B_PALPR
8	28	15.6	17	1 LCK_RAT
9	28	15.6	25	1 COX6_NEUCR
10	27.5	15.4	15	1 MK1_PALPR
11	27	15.1	15	1 MK2A_PALPR
12	27	15.1	16	1 MK3_PALPR
13	27	15.1	23	1 ECDB_LYMDI
14	27	15.1	28	1 HORC_HORSP
15	26	14.5	14	1 LPW_EGOLI
16	26	14.5	20	1 COXB_THUOB
17	26	14.5	24	1 LPA2_BACSU
18	26	14.5	28	1 PHAC_ANASP
19	25.5	14.2	21	1 ANT3_MESAU
20	25	14.0	11	1 RRPL_CHAV
21	25	14.0	15	1 NU03_SOLTU
22	25	14.0	28	1 UC05_MAIZE
23	24.5	13.7	27	1 NEUV_BUFRE
24	24	13.4	11	1 TKNA_RANRI
25	24	13.4	11	1 FKND_RANCA
26	24	13.4	16	1 TIBD_FELCA
27	24	13.4	16	1 TRVP_FELCA
28	23	12.8	10	1 UP22_HUMAN
29	23	12.8	12	1 FAR1_CALVO
30	23	12.8	13	1 TP13_PHYRO
31	23	12.8	13	1 UN02_PINPS
32	23	12.8	13	1 UP51_UPEIN
33	23	12.8	20	1 FIBB_FELCA

34 23 12.8 21 1 CFPA\_TREPH P56738 treponema p  
35 23 12.8 21 1 EPAB\_HUMAN P29322 homo sapien  
36 23 12.8 21 1 PSBF\_SYNVU P12239 synechococc  
37 23 12.8 30 1 PLMS\_SQUAC P82542 squalus aca  
38 23 12.8 31 1 EFTU\_STRLU P52390 streptomyce  
39 23 12.8 31 1 MALK\_PHOLU P41124 photorhabdu  
40 23 12.8 31 1 RECX\_METCL P37865 methylomona  
41 22 12.3 15 1 CXA2\_CONAL P56640 conus aulic  
42 22 12.3 15 1 UC16\_MAIZE P80622 zea mays (m  
43 22 12.3 16 1 MMPX\_SOLTU P80501 solanum tub  
44 22 12.3 20 1 COXF\_ONCMY P80329 oncorhynch  
45 22 12.3 25 1 PKI2\_SOLTU P24744 solanum tub

## ALIGNMENTS

### RESULT 1

NU01\_SOLTU STANDARD; PRT; 29 AA.  
AC P80267;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE NADH-UBIQUINONE OXIDOREDUCTASE 14 KDA SUBUNIT (EC 1.6.5.3)  
DE (EC 1.6.99.3) (COMPLEX I-14KD) (CI-14KD) (FRAGMENT).  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV, BINTJE; TISSUE=Tuber;  
RX MEDLINE=94124587; Pubmed=8294484;  
RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,  
RA Grohmann L.;  
RT \*Purification of the NADH:ubiquinone oxidoreductase (complex I) of  
RT the respiratory chain from the inner mitochondrial membrane of  
RT Solanum tuberosum\*;  
RL J. Biol. Chem. 269:2263-2269(1994).  
CC -I- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
CC TO BE UBIQUINONE.  
CC -I- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
CC -I- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.  
CC -I- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER  
CC MEMBRANE.  
CC PIR; I49732; I49732.  
DR Oxidoreductase; NAD; 29  
FT NON\_TER 29  
SQ SEQUENCE 29 AA; 3269 MW; E2B4DFB558D423D4 CRC64;

Query Match 22.3%; Score 40; DB 1; Length 29;  
Best Local Similarity 46.7%; Pred. No. 20;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

### QY 17 PEDIYPSNPTDDVSS 31

Db 13 PDDEFDPNNPKDVA 27

### RESULT 2

CY35\_DESAC STANDARD; PRT; 30 AA.  
ID CY35\_DESAC  
AC P81079;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CYTOCHROME C3, 50 KDA (FRAGMENT).  
OS Desulfuromonas acetoxidans (Chloropseudomonas ethylica).  
OC Bacteria; Proteobacteria; delta subdivision; Desulfuromonas group;

OC Desulfuromonas.  
 OX NCBI\_TaxID=891;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=97419123; PubMed=9271490;  
 RA Bruschi M., Woudstra M., Guigliarelli B., Asso M., Lojou E.,  
 RA Petillot Y., Abergel C.;  
 RT "Biochemical and spectroscopic characterization of two new  
 RL Biochromes isolated from Desulfuromonas acetoxidans.";  
 CC Biochemistry 36:10601-10608(1997).  
 CC -!- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH  
 CC PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME  
 CC DEHYDROGENASE TO FERREDOXIN. THE REDOX POTENTIALS OF THIS  
 CC CYTOCHROME ARE -140 MV, -210 MV AND -240 MV.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.  
 CC -!- PTM: BINDS 4 HEMES.  
 CC -!- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.  
 DR InterPro: IPR000345; CYC\_heme\_bind.  
 DR PROSITE: PS00190; CYTOCHROME\_C; PARTIAL.  
 KW Electron transport; Sulfate respiration; Heme; Periplasmic.  
 FT NON\_TER 30 30  
 SQ SEQUENCE 30 AA; 3018 MW; B0D52603DD5069B8 CRC64;

Query Match 19.0%; Score 34; DB 1; Length 30;  
 Best Local Similarity 47.1%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 RDGTRYVQKGEYRTNP 18  
 |||| ||: || |  
 Db 7 RDGTIAATGKAKTLAE 23

RESULT 3  
 FEDG\_AMEYE  
 ID FEDG\_AMEYE STANDARD; PRT; 24 AA.  
 AC P80707;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE FORMATE ESTER DEHYDROGENASE, GAMMA CHAIN (EC 1.2.99.-) (FEDH  
 DE (FRAGMENT).  
 OS Amycolatopsis methanolica.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;  
 OC Amycolatopsis.  
 OX NCBI\_TaxID=1814;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=NCIB 11946;  
 RX MEDLINE=96140591; PubMed=8554333;  
 RA Kim S.W., Luykx D.M.A.M., de Vries S., Duine J.A.;  
 RT "A second molybdoenzyme aldehyde dehydrogenase from Amycolatopsis  
 RL methanolica NCIB 11946.";  
 Arch. Biochem. Biophys. 325:1-7(1996).  
 CC -!- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA  
 CC CHAIN.  
 CC Oxidoreductase.  
 KW NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2746 MW; A93A8EA007D0FC6B CRC64;

Query Match 18.4%; Score 33; DB 1; Length 24;  
 Best Local Similarity 53.8%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 DGTRYVQKGEYRT 15  
 |||| | | | |  
 Db 8 DGTSYDEVPEPT 20

RESULT

GDO\_TRIMO  
 ID GDO\_TRIMO STANDARD; PRT; 28 AA.  
 AC P02865;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE OMEGA-GLIADIN (FRAGMENT).  
 OS Triticum monococcum (Einkorn wheat) (Small spelt).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticaceae; Triticum.  
 OX NCBI\_TaxID=4568;  
 RN [1]  
 RP SEQUENCE.  
 RA Shewry P.R., Autran J.-C., Nimmo C.C., Lew E.J.-L., Kasarda D.D.;  
 RT "N-terminal amino acid sequence homology of storage protein components  
 RL from barley and a diploid wheat.";  
 Nature 286:520-522(1980).  
 DR PIR: A03356; A03356.  
 KW Seed storage protein.  
 FT NON\_TER 28 28  
 SQ SEQUENCE 28 AA; 3343 MW; EA368C84E893FAA7 CRC64;

Query Match 18.4%; Score 33; DB 1; Length 28;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 16 NPEDIYPSNP 25  
 ||: || |  
 Db 13 SPQQLYPOQP 22

RESULT 5  
 PYRR\_PYRAP  
 ID PYRR\_PYRAP STANDARD; PRT; 20 AA.  
 AC P37362; P80307;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PYRRHOCORICIN.  
 OS Pyrrhocoris apterus (Sap sucking bug).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;  
 OC Heteroptera; Pentheteroptera; Pentatomomorpha; Lygaeoidea;  
 OC Pyrrhocoridae; Pyrrhocoris.  
 OX NCBI\_TaxID=37000;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Hemolymph;  
 RX MEDLINE=94271176; PubMed=8002963;  
 RA Cociancich S., Dupont A., Hegy G., Lanot R., Holder F., Hetru C.,  
 RA Hoffmann J.A., Bulet P.;  
 RT "Novel inducible antibacterial peptides from a hemipteran insect, the  
 RL sap-sucking bug Pyrrhocoris apterus.";  
 Biochem. J. 300:567-575(1994).  
 RN [2]  
 RP CARBOHYDRATE-LINKAGE SITE THR-11.  
 RX MEDLINE=99177428; PubMed=10076062;  
 RA Hoffmann R., Bulet P., Urge L., Otvoes L. Jr.;  
 RT "Range of activity and metabolic stability of synthetic antibacterial  
 RL glycopeptides from insects.";  
 Biochim. Biophys. Acta 1426:459-467(1999).  
 CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE. AFFECTS GRAM-NEGATIVE BACTERIA  
 CC E. COLI 1106, P. AERUGINOSA, E. COLI D22 AND E. CLOACAE AND  
 CC GRAM-POSITIVE BACTERIA M. LUTEUS AND B. SUBTILIS.  
 CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-  
 CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.  
 DR PIR: S44465; S44465.  
 KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.  
 FT CARBOHYD 11 11  
 SQ SEQUENCE 20 AA; 2341 MW; F4320EC2FF29462C CRC64;

```

CC -!- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 16 AA; 2040 MW; AF21407D063B9462 CRC64;

Query Match 16.5%; Score 29.5; DB 1; Length 16;
Best Local Similarity 41.2%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

QY 8 VOKGEY---RTNPEDIYPSN 24
Db 1 VDKGSLPRPTPPRIYRN 20

RESULT 6
PSAF_SYNPF STANDARD; PRT; 29 AA.
AC P31083;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT III (PSI-F) (FRAGMENT).
GN PSF.
OS Synchococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE.
RX MEDLINE=91355213; PubMed=1653017;
RA Li N., Warren P.V., Golbeck J.H., Frank G., Zuber H., Bryant D.A.;
RT "Polypeptide composition of the Photosystem I complex and the
RT Photosystem I core protein from Synchococcus sp. PCC 6301.";
RL Blochim. Biophys. Acta 1059:215-225(1991).
CC -!- FUNCTION: PROBABLY PARTICIPATES IN EFFICIENCY OF ELECTRON TRANSFER
CC FROM PLASTOCYANIN TO P700 (OR CYTOCHROME C553 IN ALGAE AND
CC CYANOBACTERIA). THIS PLASTOCYANIN-DOCKING PROTEIN CONTRIBUTES
CC TO THE SPECIFIC ASSOCIATION OF PLASTOCYANIN TO PS I.
CC -!- SIMILARITY: BELONGS TO THE PSAF FAMILY.
KW Photosynthesis; Photosystem I.
FT NON_TER 29
SQ SEQUENCE 29 AA; 3040 MW; 535DDC7A63B18711 CRC64;

Query Match 16.8%; Score 30; DB 1; Length 29;
Best Local Similarity 25.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 DQTRYVQKGEYRTNPE 18
Db 10 ESPRFIQRAEAAATPQ 25

RESULT 7
MK2B_PALPR STANDARD; PRT; 16 AA.
AC P80410;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN IIB.
OS Palomona prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomidae;
OC Pentatomidae; Palomona.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomona prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.

```

QY 23 SNPTDD 28  
 DB 6 SNPEDD 11

## RESULT 9

COX6\_NEUCR STANDARD; PRT; 25 AA.  
 AC Q01359;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE VI (EC 1.9.3.1) (FRAGMENT).  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86085927; PubMed=3001085;  
 RA Sachs M.S., David M., Werner S., Rajbhandary U.L.;  
 RT "Nuclear genes for cytochrome c oxidase subunits of Neurospora  
 crassa. Isolation and characterization of cDNA clones for subunits  
 IV, V, VI, and possibly VII.";  
 RL J. Biol. Chem. 261:869-873(1986).  
 CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C  
 CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -!- SUBUNIT: COMPOSED OF AT LEAST 11 SUBUNITS.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M12118; AAA33576.1;  
 KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.  
 FT NON\_TER 1  
 SQ SEQUENCE 25 AA; 2870 MW; 4D8206C608122EEE CRC64;

Query Match 15.6%; Score 28; DB 1; Length 25;  
 Best Local Similarity 62.5%; Pred. No. 7.6e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TNPEDYIP 22  
 DB 14 TLKEDLYP 21

## RESULT 10

MK1\_PALPR STANDARD; PRT; 15 AA.  
 AC P80408;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE METALNIKOWIN I.  
 OS Palomena prasina.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;  
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomolidae;  
 OC Pentatomidae; Palomena.  
 OX NCBI\_TaxID=55431;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Hemolymph;

QY 15 TNPEDYIP 22  
 DB 14 TLKEDLYP 21

RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;  
 RT "The inducible antibacterial peptides of the hemipteran insect  
 RT Palomena prasina. Identification of a unique family of proline-rich  
 RT peptides and of a novel insect defensin.";  
 RL J. Insect Physiol. 42:81-89(1996).  
 CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE  
 CC BACTERIA.  
 CC -!- INDUCTION: BY BACTERIAL INFECTION.  
 KW Antibiotic; Insect immunity.  
 SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

Query Match 15.4%; Score 27.5; DB 1; Length 15;  
 Best Local Similarity 41.2%; Pred. No. 5.1e+02;  
 Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 8 VQGEYRTNPEDIYPSN 24  
 DB 1 VDKPDYRPRP ---PPN 14

## RESULT 11

MK2A\_PALPR STANDARD; PRT; 15 AA.  
 AC P80409;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE METALNIKOWIN IIA.  
 OS Palomena prasina.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;  
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomolidae;  
 OC Pentatomidae; Palomena.  
 OX NCBI\_TaxID=55431;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Hemolymph;  
 RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;  
 RT "The inducible antibacterial peptides of the hemipteran insect  
 RT Palomena prasina. Identification of a unique family of proline-rich  
 RT peptides and of a novel insect defensin.";  
 RL J. Insect Physiol. 42:81-89(1996).  
 CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE  
 CC BACTERIA.  
 CC -!- INDUCTION: BY BACTERIAL INFECTION.  
 KW Antibiotic; Insect immunity.  
 SQ SEQUENCE 15 AA; 1893 MW; 23835D063B946299 CRC64;

Query Match 15.1%; Score 27; DB 1; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 VQGEYRTNP 17  
 DB 1 VDKPDYRPRP 10

## RESULT 12

MK3\_PALPR STANDARD; PRT; 16 AA.  
 AC P80411;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE METALNIKOWIN III.  
 OS Palomena prasina.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;  
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomolidae;  
 OC Pentatomidae; Palomena.  
 OX NCBI\_TaxID=55431;

```
RN SEQUENCE.
RP TISSUE=Hemolymph;
RC Chernysh S., Cocciandich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RL Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin."
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
CC Antibiotic; Insect immunity.
KW SEQUENCE 16 AA; 2024 MW; A9E3835D063B9462 CRC64;
SQ

Query Match 15.1%; Score 27; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 VOKGEYRTNP 17
DB 1 VDKPDYRPRP 10

RESULT 13
ECDB_LYMDI STANDARD; PRT; 23 AA.
ID ECDB_LYMDI STANDARD; PRT; 23 AA.
AC P80939;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE TESTIS ECDYSIOTROPIN PEPTIDE B (TB).
OS Lymantria dispar (Gypsy moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Lymantriidae; Lymantria.
OX NCBI_TaxID=13123;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=97387807; PubMed=9243792;
RA Loeb M.J., Wagner R.M., Woods C.W., Gelman D.G., Harrison D.,
RA Bell R.A.;
RT "Naturally occurring analogs of Lymantria testis ecdysiotropin, a
RT gonadotropin isolated from brains of Lymantria dispar pupae."
RL Arch. Insect Biochem. Physiol. 36:37-50(1997).
CC -!- FUNCTION: STIMULATES SYNTHESIS OF ECDYSTEROID IN THE TESTES
CC OF LARVAE AND PUPAE.
SQ SEQUENCE 23 AA; 2550 MW; 21AF7C84232032A6 CRC64;

Query Match 15.1%; Score 27; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 18 EDIYPSNPTD 27
DB 6 EDVEPLNAD 15

RESULT 14
HOC_HORSP STANDARD; PRT; 28 AA.
ID HOC_HORSP STANDARD; PRT; 28 AA.
AC P02864;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE C-HORDIN (FRAGMENT).
OS Hordeum spontaneum (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=77009;
```

```
RN SEQUENCE.
RP Shewry P.R., Autran J.-C., Nimmo C.C., Lew E.J.-L., Kasarda D.D.;
RT "N-terminal amino acid sequence homology of storage protein components
RL from barley and a diploid wheat."
RL Nature 286:520-522(1980).
CC -!- FUNCTION: SULFUR-POOR SEED STORAGE PROTEIN.
CC -!- TISSUE SPECIFICITY: DEVELOPING ENDOSPERM.
CC PIR: A03355; A03355.
KW Seed storage protein; Multigene family.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3333 MW; 8DBA2DF2494775AA CRC64;

Query Match 15.1%; Score 27; DB 1; Length 28;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 YPSNP 25
DB 22 YQNP 26

RESULT 15
LPW_ECOLI STANDARD; PRT; 14 AA.
ID LPW_ECOLI STANDARD; PRT; 14 AA.
AC P03053;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRP OPERON LEADER PEPTIDE.
GN TRPL OR TRPE OR B1265 OR Z2545 OR ECS1837.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82150258; PubMed=7038627;
RA Yanofsky C., Platt T., Crawford I.P., Nichols B.P., Christie G.E.,
RA Horowitz H., van Cleemput M., Wu A.M.;
RT "The complete nucleotide sequence of the tryptophan operon of
RT Escherichia coli."
RL Nucleic Acids Res. 9:6647-6668(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=76240562; PubMed=781271;
RA Squires C., Lee F., Bertrand K., Squires C.L., Bronson M.J.,
RA Yanofsky C.;
RT "Nucleotide sequence of the 5' end of tryptophan messenger RNA of
RT Escherichia coli."
RL J. Mol. Biol. 103:351-381(1976).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80101455; PubMed=118451;
RA Oxender D.L., Zurawski G., Yanofsky C.;
RT "Attenuation in the Escherichia coli tryptophan operon: role of RNA
RT secondary structure involving the tryptophan codon region."
RL Proc. Natl. Acad. Sci. U.S.A. 76:5524-5528(1979).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MGI1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1233-1238(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / EDL933 / ATCC 700927;
```

RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
O157:H7 and genomic comparison with a laboratory strain K-12";  
RL DNA Res. 8:11-22(2001).  
CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS  
CC OF TRYPTOPHAN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J01714; AAA57296.1; -  
DR EMBL; A04494; CAA00361.1; -  
DR EMBL; AE000224; AAC74347.1; -  
DR EMBL; AE005380; AAG56550.1; -  
DR EMBL; AP002556; BAB35260.1; -  
DR PIR; A03589; LFECW.  
DR EcoGene; EG11274; trpL.  
KW Tryptophan biosynthesis; Leader peptide; Complete proteome.  
SQ SEQUENCE 14 AA; 1723 MW; 5B79306E3E804A37 CRC64;

Query Match 14.5%; Score 26; DB 1; Length 14;  
Best Local Similarity 50.0%; Pred. No. 7.6e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 YVQKGEYRTN 16  
: | | : | |  
Db 5 FVLKGWRTS 14

Search completed: February 21, 2002, 16:37:51  
Job time: 181 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:34:00 ; Search time 22 seconds

(without alignments)  
212.760 Million cell updates/sec

Title: US-08-753-851-3

Perfect score: 179

Sequence: 1 CRDTRYVQGEYRTNPEDIYPSNPTDDVSS 32

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 13439

Minimum DB seq length: 0

Maximum DB seq length: 32

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL.17.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organella.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.vertibrate.\*  
14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	39	21.8	26	5 Q9U609	Q9U609 plasmodium
2	35	19.6	25	8 Q33292	Q33292 zea mays (m
3	35	19.6	30	11 Q9QUV6	Q9quv6 cavia (guin
4	32.5	18.2	28	9 Q38269	Q38269 bacterioph
5	32	17.9	20	5 Q9TWI5	Q9twi5 ceratitidis c
6	32	17.9	26	2 Q86139	Q86139 clostridium
7	32	17.9	28	12 Q9E1F3	Q9elt3 rabies viru
8	31.5	17.6	19	11 Q9QV20	Q9qv20 mus sp. per
9	31.5	17.6	30	1 Q9HMB0	Q9hmb0 halobacteri
10	31	17.3	20	5 Q9TWI8	Q9twi8 ceratitidis c
11	31	17.3	24	7 Q9TNS5	Q9tns5 homo sapien
12	31	17.3	28	12 Q9E1F4	Q9elt4 rabies viru
13	30.5	17.0	27	11 Q9CRF6	Q9crf6 mus musculu
14	30	16.8	21	4 Q16218	Q16218 homo sapien
15	30	16.8	26	11 Q9QIU8	Q9qui8 rattus norv
16	30	16.8	32	11 Q70125	Q70125 mus musculu
17	29.5	16.5	27	5 Q23745	Q23745 ctenodrilus
18	29	16.2	25	8 Q9T2K5	Q9t2k5 pisum sativ
19	29	16.2	30	1 Q9V2V2	Q9v2v2 methanobact

20	29	16.2	30	2 Q03618	Q03618 streptomyce
21	29	16.2	30	2 Q52184	Q52184 methanobact
22	29	16.2	31	8 Q37098	Q37098 zea mays (m
23	29	16.2	32	2 Q47227	Q47227 escherichia
24	28.5	15.9	25	10 Q9S8N5	Q9s8n5 lycopersico
25	28.5	15.9	29	6 P79302	P79302 sus scrofa
26	28.5	15.9	29	12 Q9IQI7	Q9iqi7 human immun
27	28	15.6	20	2 Q9RLP2	Q9rlp2 neisseria g
28	28	15.6	21	6 Q18764	Q18764 sus scrofa
29	28	15.6	22	1 Q9UWK5	Q9uwk5 methanobact
30	28	15.6	24	2 Q54460	Q54460 enterobacte
31	28	15.6	25	11 Q9QVH5	Q9qvhs rattus sp.
32	28	15.6	26	4 Q9NOV5	Q9nqv5 homo sapien
33	28	15.6	27	2 P71417	P71417 escherichia
34	28	15.6	28	12 Q9E1T5	Q9elt5 rabies viru
35	28	15.6	30	2 Q9R4J2	Q9r4j2 helicobacte
36	27.5	15.4	24	11 Q64234	Q64234 rattus sp.
37	27.5	15.4	27	1 Q9UWG8	Q9uwg8 thermococcu
38	27.5	15.4	29	12 Q9QAR4	Q9qar4 bovine coro
39	27.5	15.4	29	12 Q9DH48	Q9dh48 bovine coro
40	27	15.1	20	5 Q25281	Q25281 leishmania
41	27	15.1	20	6 Q28079	Q28079 bos taurus
42	27	15.1	22	2 Q9R4Q5	Q9r4q5 borrelia bu
43	27	15.1	22	9 Q37940	Q37940 lactococcus
44	27	15.1	22	12 Q84254	Q84254 bovine papi
45	27	15.1	23	13 Q9PRV6	Q9prv6 anguilla ja

## ALIGNMENTS

RESULT 1

Q9U609 PRELIMINARY; PRT; 26 AA.  
AC Q9U609;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
DE MEROZOITE SURFACE PROTEIN 1 (FRAGMENT).  
GN MSP1.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
OX NCBI\_taxid=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=94/98;  
RA Kirchgatter K., Branquinho M.S., Salles T.M., Lian Y.C., Roberto J.,  
RA Wunderlich G., Di Santi S.M.;  
RT "Nosocomial malaria transmission identified by molecular typing of  
RT Plasmodium falciparum.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF176817; AAD54320.1; -.  
KW Merozoite.  
FT NON\_TER 1 1  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 2921 MW; B26309B1A2117127 CRC64;

Query Match 21.8%; Score 39; DB 5; Length 26;

Best Local Similarity 47.4%; Pred. No. 1.1e+02;

Matches 9; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 14 RTNPEDIYPSNPTDDVSS 32

Db 4 RTNPDS----NSSDSDAKS 18

||||| | : | | |

RESULT 2

Q33292 PRELIMINARY; PRT; 25 AA.

ID Q33292

AC Q33292;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)



DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE CHLOROPLAST ATPase-BETA SUBUNIT GENE N (FRAGMENT).  
OS Zea mays (Maize).  
OC Chloroplast.  
OG Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86136006; PubMed=3004965;  
RA Bradley D., Gatenby A.A.;  
RT "Mutational analysis of the maize chloroplast ATPase-beta subunit gene  
RT promoter: the isolation of promoter mutants in E. coli and their  
RT characterization in a chloroplast in vitro transcription system.";  
RL EMBO J. 4:3641-3648(1985).  
DR EMBL; X03396; CAA27131.1; -.  
KW Chloroplast.  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 2603 MW; A987B062A134C6AE CRC64;

Query Match 19.6%; Score 35; DB 8; Length 25;  
Best Local Similarity 38.9%; Pred. NO. 3.9e+02;  
Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 14 RTNPDIYPSNPDDVVS 31

||||| : : : : :  
DB 2 RTNPPTSFGDPLESTCS 19

RESULT 3

Q3QUV6 PRELIMINARY; PRT; 30 AA.

AC Q3QUV6;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE AMINOPEPTIDASE P (FRAGMENT).

OS Cavia (guinea pigs).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.

OX NCBI\_TaxID=10140;

RN [1]

RP SEQUENCE.

RX MEDLINE=95110325; PubMed=7811266;

RA Denslow N.D., Ryan J.W., Nguyen H.P.;

RT "Guinea pig membrane-bound aminopeptidase P is a member of the proline

RT peptidase family.";

RL Biochem. Biophys. Res. Commun. 205:1790-1795(1994).

SQ SEQUENCE 30 AA; 3411 MW; A9F6A61854E44A25 CRC64;

Query Match

Best Local Similarity 19.6%; Score 35; DB 11; Length 30;

Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 7 YVQKGEYRTNPDIY 21

||||| : : : : :  
DB 10 YVQDGEFGIRIEDVF 24

RESULT 4

Q38269

ID Q38269 PRELIMINARY; PRT; 28 AA.

AC Q38269;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE ORF28.

OS Bacteriophage lambda.

OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;

OC Lambda phage group.

OX NCBI\_TaxID=10710;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=73215915; PubMed=4515613;

RA Weigelt P.H., Englund P.T., Murray K., Old R.W.;

RT "The 3'-terminal nucleotide sequences of bacteriophage lambda DNA.";

RL Proc. Natl. Acad. Sci. U.S.A. 70:1151-1155(1973).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=78234064; PubMed=354508;

RA Adhya S., Gottesman M.;

RT "Control of transcription termination.";

RL Annu. Rev. Biochem. 47:967-996(1978).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=82059489; PubMed=6458018;

RA Ineichen K., Shepherd J.C., Bickie T.A.;

RT "The DNA sequence of the phage lambda genome between PL and the gene

bet.";

RL Nucleic Acids Res. 9:4639-4653(1981).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=83189071; PubMed=6221115;

RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;

RT "Nucleotide sequence of bacteriophage lambda DNA.";

RL J. Mol. Biol. 162:729-773(1982).

RN [5]

RP SEQUENCE FROM N.A.

RA Daniels D.L., Schroeder J.L., Szybalski W., Sanger F., Blattner F.R.;

RT (in) Hendrix R.W., Roberts J.W., Stahl F.W., Weisberg R.A. (eds.);

RL Lambda II, pp.469-517, Cold Spring Harbor Laboratory Press,

RL New York (1983).

RN [6]

RP SEQUENCE FROM N.A.

RA Daniels D.L., Schroeder J.L., Szybalski W., Sanger F., Coulson A.R.,

RA Hong G.F., Hill D.F., Petersen G.B., Blattner F.R.;

RT (in) Hendrix R.W., Roberts J.W., Stahl F.W., Weisberg R.A. (eds.);

RL Lambda II, pp.519-674, Cold Spring Harbor Laboratory Press,

RL New York (1983).

RN [7]

RP SEQUENCE FROM N.A.

RX MEDLINE=84170247; PubMed=6324174;

RA Hohn B.;

RT "DNA sequences necessary for packaging of bacteriophage lambda DNA.";

RL Proc. Natl. Acad. Sci. U.S.A. 80:7456-7460(1983).

RN [8]

RP SEQUENCE FROM N.A.

RX MEDLINE=85074479; PubMed=6096022;

RA Craig N.L., Nash H.A.;

RT "E. coli integration host factor binds to specific sites in DNA.";

RL Cell 39:707-716(1984).

RN [9]

RP SEQUENCE FROM N.A.

RX MEDLINE=85083055; PubMed=6096550;

RA Edling T.D., Cooley T.E., Richards S.H., Ihler G.M.;

RT "Long range base-pairing in the leftward transcription unit of

RT bacteriophage lambda. Characterization by electron microscopy and

RT computer-aided sequence analysis.";

RL J. Mol. Biol. 179:351-365(1984).

RN [10]

RP SEQUENCE FROM N.A.

RX MEDLINE=85083085; PubMed=6096564;

RA Frackman S., Siegle D.A., Feiss M.;

RT "A functional domain of bacteriophage lambda terminase for prohead

RT binding.";

RL J. Mol. Biol. 180:283-300(1984).

RN [11]

RP SEQUENCE FROM N.A.

RX MEDLINE=85134899; PubMed=6241264;

RA Place N., Fien K., Mahoney M.E., Wulff D.L., Ho Y.S., Debouck C.,

RA Rosenberg M., Shih M.C., Gussin G.N.;

RT "Mutations that alter the DNA binding site for the bacteriophage

RT lambda cII protein and affect the translation efficiency of the cII

RT gene.";  
RL J. Mol. Biol. 180:865-880(1984).  
RN [12]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84119525; PubMed=6229793;  
RA Wulff D.L., Mahoney M., Shatzman A., Rosenberg M.;  
RT "Mutational analysis of a regulatory region in bacteriophage lambda  
RT that has overlapping signals for the initiation of transcription and  
RT translation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:555-559(1984).  
RN [13]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84222010; PubMed=6233610;  
RA Warren F., Das A.;  
RT "Formation of termination-resistant transcription complex at phage  
RT lambda nut locus: effects of altered translation and a ribosomal  
RT mutation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:3612-3616(1984).  
RN [14]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85232076; PubMed=2408965;  
RA Coleclough C., Erlitz F.L.;  
RT "Use of primer-restriction-end adapters in a novel cDNA cloning  
RT strategy.";  
RL Gene 34:305-314(1985).  
RN [15]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85142185; PubMed=3156406;  
RA Peltz S.W., Brown A.L., Hasan N., Podhajaska A.J., Szybalski W.;  
RT "Thermosensitivity of a DNA recognition site: activity of a truncated  
RT nutL antiterminator of coliphage lambda.";  
RL Science 228:91-93(1985).  
RN [16]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87280224; PubMed=3038914;  
RA Chen C.Y., Richardson J.P.;  
RT "Sequence elements essential for rho-dependent transcription  
RT termination at lambda TRL.";  
RL J. Biol. Chem. 262:11292-11299(1987).  
DR EMBL; J02459; AAA96576.1; -;  
SQ SEQUENCE 28 AA; 3699 MW; 2011F5C89C669BE2 CRC64;

Query Match 18.2%; Score 32.5; DB 9; Length 28;  
Best Local Similarity 33.3%; Pred. No. 9.7e+02;  
Matches 7; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

QY 12 EYRTNPEDI---YPSNPTDDD 29  
I: :|:|: | | | |  
Db 7 EFEEHPQDVMEQYQDYQYDYD 27

RESULT 5  
Q9TWI5 PRELIMINARY; PRT; 20 AA.  
AC Q9TWI5;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE MAJOR CUTICULAR PROTEIN (FRAGMENT).  
OS Ceratitis capitata (Mediterranean fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Tephriticoidea; Tephritidae; Ceratitis.  
OX NCBI\_TaxID=72113;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95279142; PubMed=7759285;  
RA Tsilivas A., Koromantzou E., Allen G., Dimitriadis G.;  
RT "Partial N-terminal sequences of larval cuticular proteins from the  
RT dipteran Ceratitis capitata.";  
RL Hereditas 122:79-83(1995).  
SQ SEQUENCE 20 AA; 2329 MW; C6ABB0B44ED70B88 CRC64;

Query Match 17.9%; Score 32; DB 5; Length 20;  
Best Local Similarity 42.9%; Pred. No. 7.8e+02;  
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
QY 8 VQGEYRTNPEDIY 21  
I: | | | | | | | |  
Db 6 VIRSEVNPDSIY 19

RESULT 6  
O86139 PRELIMINARY; PRT; 26 AA.  
ID O86139;  
AC O86139;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE RUBREDOXIN (FRAGMENT).  
OS Clostridium butyricum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1492;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MI;  
RA Gerard P., Amine J., Raval G., Young M., Petitdemange H.;  
RT "Distribution of the rubredoxin gene among the Clostridium butyricum  
RT species.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ010147; CAA09017.1; -;  
DR HSSP; P00268; IIRO.  
DR InterPro; IPR001052; Rubredoxin.  
DR InterPro; IPR003252; Rubredxn\_domain.  
DR Pfam; PF00301; rubredoxin; 1.  
DR ProDom; PD001610; Rubredxn\_domain; 1.  
DR NON\_TER 1  
FT NON\_TER 26  
SQ SEQUENCE 26 AA; 2883 MW; 7222E7887FBB59D0 CRC64;

Query Match 17.9%; Score 32; DB 2; Length 26;  
Best Local Similarity 58.3%; Pred. No. 1e+03;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 YRTNPEDIYPSN 24  
I: | | | | | | | |  
Db 1 YIYNPEDGDPDN 12

RESULT 7.  
Q9ELT3 PRELIMINARY; PRT; 28 AA.  
AC Q9ELT3;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE L PROTEIN (FRAGMENT).  
OS rabies virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Rhabdoviridae; Lyssavirus.  
OX NCBI\_TaxID=11292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=39481;  
RA Paez A., Duarte X., Pulido S.;  
RT "Molecular epidemiology of rabies epizootics in Colombia.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF276906; AAG24801.1; -;  
FT NON\_TER 28  
SQ SEQUENCE 28 AA; 2953 MW; 2E330EA9AB8F3C6A CRC64;

Query Match 17.6%; Score 31.5; DB 1; Length 30;  
Best Local Similarity 33.3%; Pred. No. 1.4e+03;  
Matches 8: Conservative 2; Mismatches 9; Indels 5; Gaps 1;

```

RESULT 10
Q9TWW8
ID Q9TWW8 PRELIMINARY; PRT; 20 AA.
AC Q9TWW8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE MAJOR CUTICULAR PROTEIN (FRAGMENT).
OS Ceratitis capitata (Mediterranean fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Tephritoidea; Tephritidae; Ceratitidae.
OX NCBI_TaxID=7213;
RN [1]
RP SEQUENCE.
RX MEDLINE=95279142; PubMed=759285;
RA Tsiliras A., Koronantzou E., Allen G., Dimitriadis G.;
RT "partial N-terminal sequences of larval cuticular proteins from the
RT dipteran Ceratitis capitata.";
RL Hereditas 122:79-83(1995).
SQ SEQUENCE 20 AA; 2408 MW; 2BC25E2F9EC15A80 CRC64;

```

Query Match 17.3%; Score 31; DB 5; Length 20;  
Best Local Similarity 42.9%; Pred. No. 1.1e+03;  
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

RESULT	11
Q9TNS5	
ID	PRELIMINARY; PRT; 24 AA.
Q9TNS5;	
AC	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE	HUMAN LEUCOCYTE ANTIGEN BETA CHAIN DR MOLECULE HLA-DRB1 (FRAGMENT).
OS	Homo sapiens (human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=93107319; PubMed=1469092;
RA	Weyand C.M., Hickok K.C., Hunder G.G., Goronzy J.J.;
RT	"The HLA-DRB1 locus as a genetic component in giant cell arteritis.
RT	Mapping of a disease-linked sequence motif to the antigen binding site
RL	of the HLA-DR molecule.";
RL	J. Clin. Invest. 90:2355-2361(1992).
KW	MHC.
SQ	SEQUENCE 24 AA; 3019 MW; 4A23BF7821BACC60 CRC64;

```

Query Match      17.3%; Score 31; DB 7; Length 24;
Best Local Similarity 71.4%; Pred. NO. 1.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      7 YVQKGEY 13
Db      17 YLQKGRY 23

```

```

RESULT 12
Q9E1T4
ID Q9E1T4 PRELIMINARY; PRT; 28 AA.
AC Q9E1T4
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE L PROTEIN (FRAGMENT).
OS rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=40389;
RA Paez A., Duarte X., Pulido S.;
RT "Molecular epidemiology of rabies epizootics in Colombia.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF276905; AAG24800.1; -
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 2939 MW; 2E2267A9AB8F3C6A CRC64;

Query Match 17.3%; Score 31; DB 12; Length 28;
Best Local Similarity 27.8%; Pred. No. 1.6e+03;
Matches 5; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

QY 16 NPEDIV--PSNPDDVVS 31
DB 3 DPGEVDDNPDPIDSEAA 20

RESULT 13
Q9CRF6
ID Q9CRF6 PRELIMINARY; PRT; 27 AA.
AC Q9CRF6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 2410117E06RIK PROTEIN (FRAGMENT).
GN 2410117E06RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC STEM CELLS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010771; BAB27171.1; -
DR MGD; MGI:1920928; 2410117E06RIK.
FT NON_TER 1 1

```

```

SQ SEQUENCE 27 AA; 2979 MW; B8D701C9E3DDB4E0 CRC64;

Query Match 17.0%; Score 30.5; DB 11; Length 27;
Best Local Similarity 46.7%; Pred. No. 1.8e+03;
Matches 7; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 17 PEDYPSNPDDVVS 31
DB 9 PPDL-PTNSSDILLS 22

RESULT 14
Q16218
ID Q16218 PRELIMINARY; PRT; 21 AA.
AC Q16218;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE DELETED IN COLORECTAL CANCER PROTEIN (FRAGMENT).
GN DCC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94320088; PubMed=8044801;
RA Reale M.A., Hu G., Zafar A.I., Getzenberg R.H., Levine S.M.,
RA Fearon E.R.;
RT "Expression and alternative splicing of the deleted in colorectal
RT cancer (DCC) gene in normal and malignant tissues.";
RL Cancer Res. 54:4493-4501(1994).
DR EMBL; S73204; AAB31484.2; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2337 MW; D2C3578D41548440 CRC64;

Query Match 16.8%; Score 30; DB 4; Length 21;
Best Local Similarity 62.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TNPEDIYP 22
DB 2 TDPVDYYP 9

RESULT 15
Q9QU18
ID Q9QU18 PRELIMINARY; PRT; 26 AA.
AC Q9QU18;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SODIUM CHANNEL BETA SUBUNIT (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Kreutz R.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U35177; AAD10398.1; -
DR EMBL; U35176; AAD10397.1; -
DR InterPro; IPR001873; ASC.
DR InterPro; PF00858; ASC; 1.
DR PRINTS; PR01078; AMINACHANNEL.
KW Ionic channel.
FT NON_TER 1 1
FT NON_TER 26 26

```

SQ SEQUENCE 26 AA; 3145 MW; B8E117C76575979C CRC64;

Query Match 16.8%; Score 30; DB 11; Length 26;  
 Best Local Similarity 44.4%; Pred. No. 2e+03;  
 Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 7 YVOKGEYRTNPEDIYPSN 24  
 | | | | | | | |  
 Db 10 YFOEFNYRTIES--PAN 25

Search completed: February 21, 2002, 16:37:14  
 Job time: 194 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:31:15 ; Search time 23.74 Seconds  
(without alignments)  
99.846 Million cell updates/sec

Title: US-08-753-851-3

Perfect score: 179

Sequence: 1 CRDGRVQKGEYRTNPDIIYPSNPTDDVSS 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 243022

Minimum DB seq length: 0

Maximum DB seq length: 32

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Genesec.1101.\*

1: /SIDS8/gcgdata/genesec/genesecp/AA1980.DAT.\*  
2: /SIDS8/gcgdata/genesec/genesecp/AA1981.DAT.\*  
3: /SIDS8/gcgdata/genesec/genesecp/AA1982.DAT.\*  
4: /SIDS8/gcgdata/genesec/genesecp/AA1983.DAT.\*  
5: /SIDS8/gcgdata/genesec/genesecp/AA1984.DAT.\*  
6: /SIDS8/gcgdata/genesec/genesecp/AA1985.DAT.\*  
7: /SIDS8/gcgdata/genesec/genesecp/AA1986.DAT.\*  
8: /SIDS8/gcgdata/genesec/genesecp/AA1987.DAT.\*  
9: /SIDS8/gcgdata/genesec/genesecp/AA1988.DAT.\*  
10: /SIDS8/gcgdata/genesec/genesecp/AA1989.DAT.\*  
11: /SIDS8/gcgdata/genesec/genesecp/AA1990.DAT.\*  
12: /SIDS8/gcgdata/genesec/genesecp/AA1991.DAT.\*  
13: /SIDS8/gcgdata/genesec/genesecp/AA1992.DAT.\*  
14: /SIDS8/gcgdata/genesec/genesecp/AA1993.DAT.\*  
15: /SIDS8/gcgdata/genesec/genesecp/AA1994.DAT.\*  
16: /SIDS8/gcgdata/genesec/genesecp/AA1995.DAT.\*  
17: /SIDS8/gcgdata/genesec/genesecp/AA1996.DAT.\*  
18: /SIDS8/gcgdata/genesec/genesecp/AA1997.DAT.\*  
19: /SIDS8/gcgdata/genesec/genesecp/AA1998.DAT.\*  
20: /SIDS8/gcgdata/genesec/genesecp/AA1999.DAT.\*  
21: /SIDS8/gcgdata/genesec/genesecp/AA2000.DAT.\*  
22: /SIDS8/gcgdata/genesec/genesecp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	179	100.0	32 15 AAR53475	CD44 peptide CD44-
2	111	62.0	21 15 AAR53477	CD44 peptide CD44-
3	96	53.6	30 15 AAR53481	CD44 peptide CD44-
4	81	45.3	28 20 AAY15201	Surrounding sequen
5	40	22.3	29 17 AAR86983	Timothy grass alle
6	40	22.3	30 17 AAR86980	Timothy grass alle
7	39	21.8	24 20 AAY08793	Yeast Nup2 protein
8	39	21.8	24 20 AAY08821	Yeast YNup2 protei
9	38	21.2	21 21 AAY53086	Anticoagulant pept
10	38	21.2	20 20 AAY01162	Polypeptide fragme
11	38	21.2	20 22 AAG64195	Protein sequence 2

12	38	21.2	20	22	AAB99270	Peptide sequence 2
13	37	20.7	24	15	AAR62696	Helper T cell epit
14	37	20.7	24	16	AAR82577	Pertussis toxin he
15	37	20.7	24	17	AAW05603	Pathogen derived T
16	37	20.7	24	21	AAW80060	T helper cell (Th)
17	37	20.7	24	21	AAW54543	Unidentified pepti
18	37	20.7	24	21	AAW58771	Amino acid sequenc
19	37	20.7	24	22	AAB84439	AHL peptide mimet
20	37	20.7	25	16	AAR80000	Multisubunit immun
21	37	20.7	32	21	AAW52323	BS106 synthetic po
22	36.5	20.4	21	19	AAW42092	BS106 polypeptide
23	36.5	20.4	21	20	AAW13470	GTF antigenic epit
24	35.5	19.8	16	20	AAW43350	Lolium perenne pro
25	35	19.6	20	15	AAR60736	LPI-22, peptide fr
26	35	19.6	20	16	AAR71561	AMB 1-2.1. Ambros
27	34.5	19.3	24	14	AAR54411	AMB 1-2.15. Ambro
28	34.5	19.3	24	14	AAR54414	AMB 1-2.15. Ambro
29	34.5	19.3	24	17	AAW02492	Bacteriophage T4 f
30	34.5	19.3	26	22	AAB74004	Vpr/Vpx motif-deri
31	34	19.0	12	16	AAR81762	Fragment of p53 bi
32	34	19.0	19	16	AAW13184	Human WBPI immunog
33	34	19.0	19	19	AAW57352	Peptide #8936 enco
34	34	19.0	31	22	AAW34899	proSKI-1 putative
35	33	18.4	10	21	AAW06281	Human Stat6 antige
36	33	18.4	18	20	AAW05230	Geminivirus replic
37	33	18.4	19	21	AAW28824	Streptavidin resid
38	33	18.4	21	18	AAW18505	Peptide derived fr
39	33	18.4	23	19	AAW66705	Human antithrombin
40	33	18.4	24	18	AAW14612	Adenovirus E1A pro
41	33	18.4	25	20	AAW25353	Peptide #11260 enc
42	33	18.4	25	22	AAW37223	Neurotrophic facto
43	33	18.4	26	15	AAW51725	Peptide derived fr
44	33	18.4	28	19	AAW66707	Amino acid sequenc
45	33	18.4	28	21	AAW93257	

#### ALIGNMENTS

#### RESULT 1

AAR53475  
ID AAR53475 standard; peptide: 32 AA.

XX AC AAR53475;

DT 01-DEC-1994 (first entry)

XX CD44 peptide CD44-3.

DE Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;

KW arthritis; tumor cell metastasis; autoimmune disease;

KW immunosuppressive.

XX Homo sapiens.

XX OS

XX WO94098111-A.

XX PD 11-MAY-1994.

XX 29-OCT-1993; 93WO-US10412.

XX 30-OCT-1992; 92US-0973339.

XX (UYDU-) UNIV DUKE.

XX Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;

XX WPI; 1994-167121/20.

XX Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-  
 CC mediated tissue damage such as occurs in the course of autoimmune  
 CC diseases, e.g. rheumatoid arthritis. It can also be used for  
 CC determining metastatic potential or to treat or prevent tumor cell  
 CC metastasis. This peptide corresponds to AA 150-180 of the CD44  
 CC protein sequence.  
 XX  
 SQ Sequence 32 AA;

Query Match 100.0%; Score 179; DB 15; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-19;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRDGRYVQKGEYRTNPEDIYPSNPTDDVSS 32  
 DB 1 CRDGRYVQKGEYRTNPEDIYPSNPTDDVSS 32

RESULT 2  
 AAR53477  
 ID AAR53477 standard; peptide; 21 AA.

XX AAR53477;  
 XX  
 DT 01-DEC-1994 (first entry)  
 XX  
 DE CD44 peptide CD44-5.  
 XX  
 KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
 KW arthritis; tumor cell metastasis; autoimmune disease;  
 KW immunosuppressive.

XX Homo sapiens.  
 XX WO9409811-A.  
 XX  
 PD 11-MAY-1994.  
 XX  
 PF 29-OCT-1993; 93WO-US10412.  
 XX  
 PR 30-OCT-1992; 92US-0973339.  
 XX  
 PA (UYDU-) UNIV DUKE.  
 XX  
 PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
 XX WPI: 1994-167121/20.  
 DR  
 XX Use of CD44 protein and new peptide derivs - for developing prods  
 XX PT for inflammation, immune-mediated tissue damage and tumour cell  
 XX PT metastasis

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-  
 CC mediated tissue damage such as occurs in the course of autoimmune  
 CC diseases, e.g. rheumatoid arthritis. It can also be used for  
 CC determining metastatic potential or to treat or prevent tumor cell  
 CC metastasis. This peptide corresponds to AA 161-180 of the CD44  
 CC protein sequence.  
 XX  
 SQ Sequence 21 AA;

Query Match 62.0%; Score 111; DB 15; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 YKRRPEDIYPSNPTDDVSS 32  
 DB 13 YKRRPEDIYPSNPTDDVSS 32

Db 2 yrtnpdiypsntddvss 21

RESULT 3  
 AAR53481  
 ID AAR53481 standard; peptide; 30 AA.

XX AAR53481;  
 XX  
 DT 01-DEC-1994 (first entry)  
 XX  
 DE CD44 peptide CD44-8.  
 XX  
 KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
 KW arthritis; tumor cell metastasis; autoimmune disease;  
 KW immunosuppressive.

XX Homo sapiens.  
 XX OS  
 XX WO9409811-A.  
 XX  
 PD 11-MAY-1994.  
 XX  
 PF 29-OCT-1993; 93WO-US10412.  
 XX  
 PR 30-OCT-1992; 92US-0973339.  
 XX  
 PA (UYDU-) UNIV DUKE.  
 XX  
 PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
 XX WPI: 1994-167121/20.  
 DR  
 XX Use of CD44 protein and new peptide derivs - for developing prods  
 XX PT for inflammation, immune-mediated tissue damage and tumour cell  
 XX PT metastasis

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-  
 CC mediated tissue damage such as occurs in the course of autoimmune  
 CC diseases, e.g. rheumatoid arthritis. It can also be used for  
 CC determining metastatic potential or to treat or prevent tumor cell  
 CC metastasis. This peptide corresponds to AA 150-161, 170-177 and  
 CC 189-198 of the CD44  
 CC protein sequence.  
 XX  
 SQ Sequence 30 AA;

Query Match 53.6%; Score 96; DB 15; Length 30;  
 Best Local Similarity 67.9%; Pred. No. 2.8e-07;  
 Matches 19; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 CRDGRYVQKGEYRTNPEDIYPSNPTDD 28  
 DB 1 CRDGRYVQKGEYRTNPEDIYPSNPTDD 20

RESULT 4  
 AAY15201  
 ID AAY15201 standard; Peptide; 28 AA.

XX AAY15201;  
 XX  
 DT 26-OCT-1999 (first entry)  
 XX  
 DE Surrounding sequence of potential assembly site E5-Rg, in CD44.  
 XX isoform; CD44; expression vector; exon; modification;  
 KW proteoglycan; glycosaminoglycan binding protein; rheumatoid;  
 KW arthritis; asthma; immunological disorder; assembly site.  
 XX

OS Synthetic.  
OS Homo sapiens.

PN WO9937317-A1.  
XX

PN 29-JUL-1999.  
XX

XX 21-JAN-1999; 99WO-US01411.  
XX

XX 24-JAN-1998; 98US-0072416.  
XX

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX

XX Aruffo AA, Bennett KL, Greenfield WB, Wolff EA;  
XX

XX WPI; 1999-478982/40.  
XX

XX New artificial proteoglycans useful for treating rheumatoid  
XX arthritis

XX Example 1; Page 30; 63pp; English.  
XX

XX This is part of the surrounding sequence of one of the potential  
XX assembly sites in a CD44 exon.

XX The recombinant fusion protein may be administered to a patient to  
XX enhance the biological activity of a GAG-binding protein within the body.

XX The protein also enhances wound healing and may be used to treat  
XX rheumatoid arthritis, asthma, chronic obstructive pulmonary disorder,  
XX Lupus, inflammatory bowel disease, psoriasis, osteoarthritis or  
XX HIV infection.

XX The fusion protein may be used to target proteoglycans to a given site,  
XX which causes local accumulation of GAG-binding proteins. Therefore the  
XX protein may be used as adjuvants for vaccination and in the targeting of  
XX chemokines to non-immunogenic tumour cells to enhance cellular antitumour  
XX responses. In addition, the fusion protein can also enhance the half-life  
XX of non-GAG binding growth factors.

XX Sequence 28 AA;  
SQ

Query Match 45.3%; Score 81; DB 20; Length 28;  
Best Local Similarity 100.0%; Pred. No. 4.1e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18. EDIYPSNPTDDDDVSS 32  
|||||

Db 1 ediypsnptdddvss 15

RESULT 5  
AAR86983

ID AAR86983 standard; Peptide; 29 AA.  
XX

XX AAR86983;  
AC

XX 26-JUL-1996 (first entry)  
XX

XX Timothy grass allergen Phl p I epitope clone c108.  
XX

XX Timothy; grass; allergen; Phl p I; epitope; clone c108; diagnosis;  
XX allergy; pollen; plant protein; immunoassay; detection; antibody;  
XX treatment; immunotherapy; immunodominant; IgE hapten;  
XX anaphylactic shock; hyposensitisation.

XX Phleum pratense.  
OS

XX WO9534578-A1.  
XX

XX 21-DEC-1995.  
XX

XX 14-JUN-1995; 95WO-SE00724.  
XX

XX 14-JUN-1994; 94SE-0002089.  
XX

XX (PHAA ) PHARMACIA AB.  
XX

XX Ball T, Kraft D, Laffer S, Sperr W, Susani M, Valent P;  
XX Valenta R, Vrtala S;  
XX WPI; 1996-049622/05.

XX Recombinant DNA encoding epitope(s) of timothy grass allergen Phl p I  
XX - useful in diagnosis and treatment of allergies, partic. to grass  
XX pollen

XX (PHAA ) PHARMACIA AB.  
PA

XX Ball T, Kraft D, Laffer S, Sperr W, Susani M, Valent P;  
PI Valenta R, Vrtala S;

XX WPI; 1996-049622/05.  
XX

XX Recombinant DNA encoding epitope(s) of timothy grass allergen Phl p I  
XX - useful in diagnosis and treatment of allergies, partic. to grass  
XX pollen

XX Claim 1; Page 19; 25pp; English.  
PS

XX The present sequence is the Timothy grass allergen Phl p I  
XX epitope clone c108, which was isolated by screening a random DNase  
XX I digested Phl p I E. coli expressed lambda gt11 cDNA library,  
XX with serum IgE from Phl p I allergic patients. The epitope can be  
XX used to (pref. in vitro) diagnose allergy to plant proteins,  
XX esp. grass pollen, by immunoassay detection of specific  
XX antibodies, for the treatment of pollen allergy, by passive admin.  
XX or active immunotherapy and for measuring (pref. in vitro)  
XX cellular reactions against the Phl p I epitope. The epitope is an  
XX immunodominant IgE hapten, and is therefore less likely to cause  
XX anaphylactic shock during hyposensitisation treatment.

XX Sequence 29 AA;  
SQ

Query Match 22.3%; Score 40; DB 17; Length 29;  
Best Local Similarity 41.2%; Pred. No. 46;

Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 RYVQKGERTNPEDIYP 22  
|| :| :| :| :|

Db 1 rytteggtkteaadvip 17

RESULT 6  
AAR86980

ID AAR86980 standard; Peptide; 30 AA.  
XX

XX AAR86980;  
AC

XX 26-JUL-1996 (first entry)  
XX

XX Timothy grass allergen Phl p I epitope clone c86.  
XX

XX Timothy; grass; allergen; Phl p I; epitope; clone c86; diagnosis;  
XX allergy; pollen; plant protein; immunoassay; detection; antibody;  
XX treatment; immunotherapy; immunodominant; IgE hapten;  
XX anaphylactic shock; hyposensitisation.

XX Phleum pratense.  
OS

XX WO9534578-A1.  
XX

XX 21-DEC-1995.  
XX

XX 14-JUN-1995; 95WO-SE00724.  
XX

XX 14-JUN-1994; 94SE-0002089.  
XX

XX (PHAA ) PHARMACIA AB.  
XX

XX Ball T, Kraft D, Laffer S, Sperr W, Susani M, Valent P;  
XX Valenta R, Vrtala S;  
XX WPI; 1996-049622/05.

XX Recombinant DNA encoding epitope(s) of timothy grass allergen Phl p I  
XX - useful in diagnosis and treatment of allergies, partic. to grass  
XX pollen



XX

Claim 1; Page 19; 25pp; English.

XX

The present sequence is the Timothy grass allergen Phl p I epitope clone c86, which was isolated by screening a random DNase I digested Phl p I E. coli expressed lambda gt10 cDNA library, with serum IgE from Phl p I allergic patients. The epitope can be used to (pref. in vitro) diagnose allergy to plant proteins, esp. grass pollen, by immunoassay detection of specific antibodies, for the treatment of pollen allergy, by passive admin. or active immunotherapy and for measuring (pref. in vitro) cellular reactions against the Phl p I epitope. The epitope is an immunodominant IgE hapten, and is therefore less likely to cause anaphylactic shock during hyposensitisation treatment.

XX

Sequence 30 AA;

SQ

Query Match 22.3%; Score 40; DB 17; Length 30;  
Best Local Similarity 41.2%; Pred. No. 48;  
Matches 7; Conservative 3; Mismatches. 7; Indels 0; Gaps 0;

QY

6 RYVQKGEYRTNPEDYIP 22

DB

2 RYTEGGTKTEAEDVIP 18

RESULT

7

AAY08793

ID AAY08793 standard; Protein; 24 AA.

XX

AC

AAY08793;

XX

DT

13-AUG-1999 (first entry)

XX

DE

Yeast Nup2 protein fragment 1.

XX

Activation sequence; transcription factor; murine; p163; p27; treatment; binding protein; DNA binding domain; effector gene; disease; infection; tumour; leukaemia; autoimmune disease; allergy; arthritis; inflammation; transplant rejection, graft-versus-host disease; circulatory disorder; blood clot; anaemia; hormonal disorder; CNS injury; yeast; Nup2.

XX

OS

Saccharomyces cerevisiae.

XX

XX

PN

EP926237-A2.

XX

PD

30-JUN-1999.

XX

PF

12-DEC-1998; 98EP-0123709.

XX

PR

20-DEC-1997; 97DE-1056975.

XX

PA

(HMRI ) HOECHST MARION ROUSSEL DEUT GMBH.

XX

PI

Buergin A, Eilers M, Sedlacek H;

XX

XX

DR

WPI; 1999-349238/30.

XX

New nucleic acid construct comprising promoter, transcription factor gene, activation sequence and effector gene - useful for gene therapy treatment of allergies, inflammation, transplant disorders and leukaemia

XX

PS

Disclosure; Page 35; 90pp; German.

XX

CC

This invention describes a novel nucleic acid construct comprising the following components (a) an activation sequence for the transcription of component b, (b) component b which is constructed from component b1 (a transcription factor activating domain), component b2 (murine p163 or p27 binding protein) and component b3 (a transcription factor DNA binding domain), (c) an activation sequence which is activated by binding of the expression product of component (b) and which induces

CC

transcription of component (d) and (d) an effector gene. The construct, preferably in a plasmid or viral vector, or cell can be used to treat a disease selected from infections, tumours, leukaemia, autoimmune diseases, allergies, arthritis, inflammations, transplant rejection, graft-versus-host disease, blood clotting disorders, circulatory disorders, anaemia, hormonal disorders and CNS injuries. This sequence represents a fragment of the Saccharomyces cerevisiae Nup2 protein which is used to describe the method of the invention.

Sequence 24 AA;

Query Match 21.8%; Score 39; DB 20; Length 24;  
Best Local Similarity 58.3%; Pred. No. 52;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 19 DIYPSNPTDDVV 30

: | | | |

Db 13 etydsnesddv 24

RESULT

8

AAY08821

ID AAY08821 standard; Protein; 24 AA.

XX

AC

AAY08821;

XX

DT

13-AUG-1999 (first entry)

XX

DE

Yeast YNup2 protein fragment 1.

XX

KW

p163; murine; p27 inhibitor; p27-induced inhibition; cell proliferation;

XX

KW

p27 binding domain; Ran binding domain; detection; screening; malignancy;

XX

KW

tumour; yeast; Nup2.

XX

OS

Saccharomyces cerevisiae.

XX

XX

PN

EP926236-A1.

XX

PD

30-JUN-1999.

XX

PF

12-DEC-1998; 98EP-0123708.

XX

XX

PR

20-DEC-1997; 97DE-1056975.

XX

PA

(HMRI ) HOECHST MARION ROUSSEL DEUT GMBH.

XX

PI

Buergin A, Eilers M, Sedlacek H;

XX

XX

DR

WPI; 1999-349237/30.

XX

PT

New p27-inhibiting protein p163 and DNA - useful for detection and/or quantification of p163 mRNA

XX

PS

Disclosure; Page 14; 68pp; German.

XX

CC

This invention describes (1) a protein that inhibits p27 and thereby arrests p27-induced inhibition of cell proliferation, (2) the protein of (1) comprising at least part of the amino acid sequence of murine p163. (3) a protein that can be derived from the protein of (2) by deletion of the p27 binding domain or the Ran binding domain, (4) a protein that can be derived from the protein of (2) by deletion of all amino acid sequences other than the p27 binding domain, (5) a protein that is the human or other mammalian species homologue of a protein as in (1)-(4), (6) DNA encoding the p163 protein, (7) antibodies and antibody fragments that bind to the p27 binding domain of a protein as above, (8) antibodies and antibody fragments that bind to the Ran binding domain of a protein as above, (9) antisense nucleic acids complementary to portions of the DNA of (6) between codons 121 and 467, (10) a nucleic acid construct coding for an antisense nucleic acid as in (9) for inhibiting the proliferation of a cell, in which DNA coding for the antisense nucleic acid sequence is linked to at least one activation sequence and is introduced into the target cell as naked DNA or as an insert in a

CC nonviral or viral vector and (11) a nucleic acid construct containing the  
 CC DNA of (6) linked to an activation sequence that permits expression of a  
 CC protein as above in a cell. The DNA of (6) can be used for detection  
 CC and/or quantification of p163 mRNA in cells and/or tissues, preferably by  
 CC Northern blotting, PCR or fluorescent in-situ hybridisation. The proteins  
 CC can be used to produce antibodies, which can be used to detect the  
 CC corresponding protein in cells, tissues or body fluids. The antisense  
 CC nucleic acids can be used to inhibit cell proliferation in vitro or in  
 CC vivo. The proteins can be used to screen for substances that inhibit the  
 CC interaction between the proteins and their cellular binding partners,  
 CC preferably using a two hybrid system or using an affinity system in which  
 CC p163 or its p27-binding domain is immobilised on a solid phase, the solid  
 CC phase is incubated with a test substance, and the inhibition of the  
 CC binding of a labelled binding partner of p163 (especially p27 or Ran) is  
 CC measured. Assays for p163 can be used to assess the malignancy of  
 CC tumours.

XX  
 XX  
 SQ Sequence 24 AA;

Query Match 21.8%; Score 39; DB 20; Length 24;  
 Best Local Similarity 58.3%; Pred. No. 52;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 19 DIYPSNPTDDV 30  
 : : ||| :|||  
 Db 13 etydsnesddv 24

RESULT 9

AAAY53086  
 ID AAAY53086 standard; peptide; 12 AA.

XX AC  
 XX AAAY53086;

DT 03-APR-2000 (first entry)

XX Anticoagulant peptide derivative #25.

DE Anticoagulant; thrombin recognition site.

XX  
 XX Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 12  
 FT /note= "amidated"

XX US6005071-A.

XX 21-DEC-1999.

XX 12-FEB-1998; 98US-0022738.

XX 28-MAR-1995; 95US-0412356.

XX 12-JAN-1988; 88US-0138611.

XX 01-AUG-1989; 89US-0388725.

XX 26-JUL-1990; 90US-0559438.

XX 11-OCT-1991; 91US-0774126.

XX 10-DEC-1992; 92US-0989998.

XX 23-SEP-1993; 93US-0126441.

XX 23-JAN-1987; 87US-0006417.

XX 21-MAY-1987; 87US-0053162.

XX (RICH ) MERRELL PHARM INC.

XX Krstenansky JL, Mao SJT;  
 PI WPI; 2000-105099/09.

XX New peptide derivatives used as anticoagulants -  
 XX Claim 27; Column 19; 11pp; English.

CC AAAY53062 to AAAY53093 represent specifically claimed anticoagulant  
 CC peptide derivatives from the present invention. AAAY53094 to AAAY53125  
 CC represent peptide derivatives given in examples from the present  
 CC invention. The peptide derivatives have a thrombin recognition site  
 CC binder mechanism of action. The peptide derivatives can be used as  
 CC anticoagulants. They can bind to the thrombin recognition site without  
 CC binding to the cleavage site of thrombin.

SQ Sequence 12 AA;

Query Match 21.2%; Score 38; DB 21; Length 12;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 11 GEYRTNPEDIYP 22  
 : : ||| ||  
 Db 1 gdfepipedayp 12

RESULT 10

AAAY01162

ID AAAY01162 standard; Protein; 20 AA.

XX AC  
 XX AAAY01162;

XX 18-MAY-1999 (first entry)

XX Polypeptide fragment encoded by gene 2.

XX Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
 KW tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
 KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
 KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis.

XX Homo sapiens.

XX WO9901020-A2.

XX 14-JAN-1999.

XX 30-JUN-1998; 98WO-US13608.

XX 12-SEP-1997; 97US-0058663.

XX 01-JUL-1997; 97US-0051381.

XX 01-JUL-1997; 97US-0051480.

XX 12-SEP-1997; 97US-0058598.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;  
 PI WPI; 1999-105683/09.

XX N-PSDB; AAX22112.

XX New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, immune deficiency diseases or blood  
 PT disorders

XX Disclosure; Page 7; 179pp; English.

XX The invention relates to nucleic acid sequences (AAX22111 to AAX22134)  
 CC encoding human secreted proteins (AAAY01135 to AAAY01158). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit number  
 CC ATCC 209118. Host cells comprising recombinant vectors containing the  
 CC nucleic acid sequences are used for the recombinant production of the  
 CC secreted proteins. The polynucleotide and amino acid sequences are useful  
 CC for are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Pathological conditions can

CC be also diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC polynucleotides, based on which tissues they are most highly expressed  
 CC in, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC polypeptide fragment encoded by a gene of the invention (see descriptor  
 CC line for gene number).

XX Sequence 20 AA;

Query Match 21.2%; Score 38; DB 20; Length 20;  
 Best Local Similarity 37.5%; Pred. No. 59;  
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 DCTRYVQKGEYRTNPE 18  
 | : | | : | | |  
 Db 2 dpkktlqmgsgfrlnpd 17

RESULT 11

AAG64195  
 ID AAG64195 standard; Peptide; 20 AA.

XX AAG64195;

XX 22-OCT-2001 (first entry)

XX Protein sequence 2.

XX Mycoplasma genitalium; complementary peptide; ligand;  
 KW protein-protein interaction; drug design.

XX Unidentified.

XX WO200142278-A2.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB04778.

XX 13-DEC-1999; 99GB-0029466.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-514238/56.

XX Complementary peptide ligands as reagents and drugs for drug discovery  
 PT programs and as lead ligands to facilitate drug design and development,  
 PT are generated from microbial genome sequences -

XX Disclosure; Page 12; 161pp; English.

XX The invention relates to a set of complementary peptide ligands  
 CC generated from Mycoplasma genitalium genome sequences. These  
 CC specific complementary peptides interact with their relevant target  
 CC proteins encoded by the microbial genome. They are capable of  
 CC antagonising or agonising specific interaction of a protein with  
 CC another protein or receptor and are thus useful as reagents and drugs,  
 CC and as lead ligands to facilitate drug design and development. They  
 CC are useful as tools for functional genomic studies, reagents for the  
 CC configuration of high-throughput screens, as a starting point for

CC medicinal chemistry manipulation, for peptide mimetics and as  
 CC therapeutic agents. The analysis and acquisition of peptide sequences  
 CC facilitates understanding of protein-protein interactions. The method  
 CC allows for analysis of an entire database at a time, thus overcoming  
 CC sampling problems. The present sequence is provided as an example  
 CC protein sequence in a description of the invention.

XX Sequence 20 AA;

Query Match 21.2%; Score 38; DB 22; Length 20;  
 Best Local Similarity 47.4%; Pred. No. 59;  
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 11 GEYRTNPEDIYPSNPTDDDD 29  
 | : | | : | | |  
 Db 1 gtfrtsredstysgdtfd 19

RESULT 12

AAB99270  
 ID AAB99270 standard; Peptide; 20 AA.

XX AAB99270;

XX 11-SEP-2001 (first entry)

XX Peptide sequence 2.

XX Plant; peptide pesticide; peptide herbicide; agricultural research.

XX Unidentified.

XX WO200142279-A2.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB04781.

XX 13-DEC-1999; 99GB-0029469.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-381629/40.

XX A set of peptide ligands for agricultural research and development of  
 PT therapeutic agents comprise specific complementary peptides to proteins  
 PT encoded by genes of plant genomes -

XX Disclosure; Page 10; 201pp; English.

XX The present invention relates to a set of peptide ligands consisting of  
 CC specific complementary peptides to proteins encoded by genes of plant  
 CC genomes e.g. from Arabidopsis thaliana (see AAG83361-AAG84504). The  
 CC peptides of the present invention are useful in an assay to identify a  
 CC peptide, especially a peptide pesticide or herbicide. The peptides are  
 CC also useful for tools for agricultural research and development. The  
 CC present peptide was used to illustrate the present invention.

XX Sequence 20 AA;

Query Match 21.2%; Score 38; DB 22; Length 20;  
 Best Local Similarity 47.4%; Pred. No. 59;  
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 11 GEYRTNPEDIYPSNPTDDDD 29  
 | : | | : | | |  
 Db 1 gtfrtsredstysgdtfd 19

## RESULT 13

AA062696  
ID AAR62696 standard; peptide; 24 AA.

AC AAR62696;

DT 10-SEP-1995 (first entry)

DE Helper T cell epitope for use in universal immune stimulator.

XX Helper T cell epitope; universal immune stimulator; invasin; hapten;

KW vaccine; pertussis toxin.

XX Bordetella pertussis.

OS WO9425060-A.

PN 10-NOV-1994.

PD 28-APR-1994; 94WO-US04832.

XX 27-APR-1993; 93US-0057166.

PR 14-APR-1994; 94US-0229275.

XX (LADD//) LADD A E.

PA (WAMB//) WAMB C Y.

PA Ladd AE, Wang CY, Zamb T;

PI WPI; 1994-357910/44.

DR Immunogenic luteinizing hormone releasing hormone peptide(s) -

XX that suppress LHRH activity in males and females

PT Claim 7; Page 25; 213pp; English.

XX Synthetic immunogenic peptides are provided in which a universal immune

CC stimulator is linked to a peptide or protein hapten containing B cell

CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes

CC potent immune responses to the coupled peptide or protein. The

CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)

CC which elicits an immune response to the coupled peptide in members of

CC a heterogeneous population expressing diverse HLA phenotypes, and (B)

CC an adjuvant peptide sequence from the invasin protein of Yersinia.

CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the

CC invasin and Th domains and between the immune stimulator and hapten

CC components. When the hapten is LHRH, then optionally the invasin domain

CC can be omitted from the immune stimulator component.

CC The present sequence represents a pertussis toxin helper T cell epitope

CC which can be used as Th in the immune stimulator.

XX Sequence 24 AA;

SQ Query Match 20.7%; Score 37; DB 15; Length 24;

Best Local Similarity 44.4%; Pred. No. 1e+02;

Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Oy 1 CRDGTTRYVQKGEYRTNPE 18

Db 6 cpngtraltvaelrgnae 23

RESULT 14

AA082577

ID AAR82577 standard; peptide; 24 AA.

XX AAR82577;

AC AAR82577;

DT 13-JUN-1996 (first entry)

XX Pertussis toxin helper T cell epitope, PT2.

DE

XX

KW

XX

OS

PN

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope; vaccine; allergy; antibody; constant heavy chain.

Bordetella pertussis.

WO9526365-A1.

05-OCT-1995.

24-MAR-1995; 95WO-US03741.

25-OCT-1994; 94US-0328912.

28-MAR-1994; 94US-0218461.

(UNBI-) UNITED BIOMEDICAL INC.

Wang CY;

WPI; 1995-351297/45.

Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper

T cell epitope - useful for eliciting antibody prodn. for allergy

treatment

Claim 3; Page 58; 87pp; English.

AAR82571-91 are helper T cell epitopes which can be used in the

preparation of a peptide immunogen that is useful in vaccines for

treating allergic reactions. In the immunogen an IgE CH4 peptide

is attached C-terminally to a series of amino acids including a

helper T cell epitope. The immunogen may also opt. contain a fatty

acid or fatty acid derivative, an invasin domain or alpha-NH2. The

immunogen produces high titres of antibodies to the effector site

in human IgE heavy chain (the CH4 domain peptide) which inhibit mast

cell activation and reduce allergen-induced IgE prodn. The immunogens

may be used in either a radially branching multimeric form or a

linearly arranged monomeric form.

Sequence 24 AA;

Query Match 20.7%; Score 37; DB 16; Length 24;

Best Local Similarity 44.4%; Pred. No. 1e+02;

Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Oy 1 CRDGTTRYVQKGEYRTNPE 18

Db 6 cpngtraltvaelrgnae 23

RESULT 15

AA05603

ID AAW05603 standard; peptide; 24 AA.

XX AAW05603;

AC AAW05603;

DT 10-DEC-1996 (first entry)

XX Pertussis toxin helper T cell epitope #2.

XX Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;

membrane anchoring domain; helper T cell; surface antigen; core antigen;

pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;

Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;

diphtheria toxin; plasmodium falciparum; circumsporozoite; E. coli Trst;

schistosoma mansoni; triose phosphate isomerase; allergenic reaction;

allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;

antihistamine; decongestant; beta-2 agonist; immunosuppression;

corticosteroid.

Hemophilus pertussis.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

PN WO9612740-A1.
XX
PD 02-MAY-1996.
XX
XX 25-OCT-1995; 95WO-US13841.
XX
PR 25-OCT-1994; 94US-0328519.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Walfield AM, Wang CY;
XX
DR WPI; 1996-230555/23.
XX
PT Peptide immunogen useful in treatment of allergy - comprises
PT membrane-bound IgE epsilon-chain peptide synthesised linearly in
PT tandem with T helper epitope peptide
XX
PS Claim 2; Page 18; 53pp; English.
XX
XX AAW05957-W05616 represent helper T cell epitopes used in the peptide
CC immunogens of the invention. This sequence represents the pertussis
CC toxin helper T cell antigen. The peptides of the invention contain one
CC of these sequences, and a membrane-bound immunoglobulin E (IgE) fragment
CC (see AAW05595 and AAW05596). The peptide immunogens of the invention can
CC be used in vaccines for the immunotherapeutic treatment of allergenic
CC reactions, including allergic rhinitis, food allergies, anaphylaxis, or
CC virally-induced asthma. The immunogens overcome the short effective
CC period of antihistamines, decongestants, and beta-2 agonists, while
CC preventing the broad immunosuppression of corticosteroids. The peptides
CC do not have the potential side effects of restlessness or sedation
CC (associated with antihistamines), associated increased morbidity in
CC asthmatics (as seen with beta-2 agonists) and adverse hormonal
CC activities (observed in corticosteroid users).
XX
SQ Sequence 24 AA;

```

Query Match 20.7%; Score 37; DB 17; Length 24;  
 Best Local Similarity 44.4%; Pred. No. 1e+02;  
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 CRDGTTRYQKGEYRTNPE 18
   | : | | | : | | | |
Db 6 cpngtraltvaelrgnae 23

```

Search completed: February 21, 2002, 16:34:28  
 Job time: 193 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:32:55 ; Search time 12.48 seconds  
(without alignments)  
57.701 Million cell updates/sec

Title: US-08-753-851-3  
Perfect score: 179  
Sequence: 1 CRDGRVQKGEYRTNPDIYSPNPTDDVSS 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 135027

Minimum DB seq length: 0  
Maximum DB seq length: 32

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	179	100.0	32	2	US-08-143-311B-3
2	111	62.0	21	2	US-08-143-311B-5
3	96	53.6	30	2	US-08-143-311B-9
4	40	22.3	29	3	US-08-750-419A-30
5	40	22.3	30	3	US-08-750-419A-28
6	39	21.8	24	4	US-09-215-221-11
7	37	20.7	24	1	US-08-446-682-8
8	37	20.7	24	2	US-08-488-351A-8
9	37	20.7	24	3	US-09-100-409A-44
10	37	20.7	24	5	PCT-US95-13841-11
11	37	20.7	32	2	US-08-563-892A-10
12	35	19.6	12	4	US-08-938-595-5
13	35	19.6	12	4	US-08-727-153-5
14	35	19.6	20	1	US-08-440-861-52
15	34	19.0	18	1	US-07-729-099-15
16	34	19.0	18	1	US-08-257-392-15
17	34	19.0	18	3	US-08-770-035-15
18	34	19.0	18	3	US-08-804-439A-113
19	34	19.0	19	1	US-08-399-696-84
20	33	18.4	23	2	US-09-017-205-83
21	33	18.4	26	1	US-08-168-091A-24
22	33	18.4	28	2	US-09-017-205-85
23	32.5	18.2	26	2	US-08-482-142-22
24	32.5	18.2	26	2	US-08-478-572-22
25	32.5	18.2	26	4	US-08-484-296-22
26	32.5	18.2	30	2	US-08-473-475A-11
27	32	17.9	25	1	US-08-264-030-11

28 31.5 17.6 27 3 US-08-744-138-9 Sequence 9, Appli  
29 31.5 17.6 27 4 US-09-241-376-9 Sequence 9, Appli  
30 31 17.3 9 2 US-08-318-856A-53 Sequence 53, Appli  
31 31 17.3 9 3 US-08-159-339A-628 Sequence 628, App  
32 31 17.3 19 1 US-08-520-977A-8 Sequence 8, Appli  
33 31 17.3 20 1 US-07-971-096-9 Sequence 9, Appli  
34 31 17.3 20 1 US-08-175-096-9 Sequence 9, Appli  
35 31 17.3 24 4 US-09-177-249-203 Sequence 203, App  
36 31 17.3 25 1 US-08-502-989-8 Sequence 8, Appli  
37 31 17.3 25 5 PCT-US94-05355-8 Sequence 8, Appli  
38 31 17.3 28 4 US-09-004-731-73 Sequence 73, Appli  
39 31 17.3 28 4 US-08-749-699-73 Sequence 73, Appli  
40 31 17.3 30 6 5240706-8 Patent No. 5240706  
41 31 17.3 30 6 5252328-4 Patent No. 5252328  
42 30 16.8 8 4 US-08-938-595-4 Sequence 4, Appli  
43 30 16.8 8 4 US-08-727-153-4 Sequence 4, Appli  
44 30 16.8 13 2 US-08-883-070-5 Sequence 5, Appli  
45 30 16.8 13 4 US-09-258-754-186 Sequence 186, App

## ALIGNMENTS

RESULT 1  
US-08-143-311B-3  
; Sequence 3, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-3

Query Match 100.0%; Score 179; DB 2; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1e-19;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRDGRYVQKGEYRTPNEDIYPSNPTDDVSS 32  
|||||  
DB 1 CRDGRYVQKGEYRTPNEDIYPSNPTDDVSS 32

## RESULT 2

US-08-143-311B-5  
; Sequence 5, Application US/08143311B  
; Patent No. 5863540

## GENERAL INFORMATION:

APPLICANT: HAYNES, BARTON F.  
APPLICANT: HALE, LAURA P.  
APPLICANT: PATTON, KAREN L.  
APPLICANT: TELEN, MARILYN J.  
APPLICANT: LIAO, HUA-XIN  
TITLE OF INVENTION: AN ADHESION MOLECULE  
NUMBER OF SEQUENCES: 28

## CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/143,311B  
FILING DATE: 29-OCT-1993

## CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,339  
FILING DATE: 30-OCT-1992  
CLASSIFICATION: 436  
APPLICATION NUMBER: 07/669,730  
FILING DATE: 15-MAR-1991  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-50

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

## US-08-143-311B-5

Query Match 62.0%; Score 111; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 5.3e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 YRTNPEDIYPSNPTDDVSS 32

|||||

DB 2 YRTNPEDIYPSNPTDDVSS 21

## RESULT 3

US-08-143-311B-5

## US-08-143-311B-9

; Sequence 9, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
APPLICANT: HAYNES, BARTON F.  
APPLICANT: HALE, LAURA P.  
APPLICANT: PATTON, KAREN L.  
APPLICANT: TELEN, MARILYN J.  
APPLICANT: LIAO, HUA-XIN  
TITLE OF INVENTION: AN ADHESION MOLECULE  
NUMBER OF SEQUENCES: 28

## CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/143,311B  
FILING DATE: 29-OCT-1993

## CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/973,339

FILING DATE: 30-OCT-1992

CLASSIFICATION: 436

APPLICATION NUMBER: 07/669,730

FILING DATE: 15-MAR-1991

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: WILSON, MARY J.

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 1579-50

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

## US-08-143-311B-9

Query Match 53.6%; Score 96; DB 2; Length 30;  
Best Local Similarity 67.9%; Pred. No. 1.3e-07;  
Matches 19; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 CRDGRYVQKGEYRTPNEDIYPSNPTDD 28

|||||

DB 1 CRDGRYVQKGEYRTPNEDIYPSNPTDD 20

## RESULT 4

US-08-750-419A-30

; Sequence 30, Application US/08750419A  
; Patent No. 6008340

## GENERAL INFORMATION:

APPLICANT: BALL, TANJA

APPLICANT: VITALA, SUSANNE

APPLICANT: SPERR, WOLFGANG

APPLICANT: VALENT, PETER

APPLICANT: SUSANI, MARKUS

APPLICANT: KRAFT, DIETRICH

APPLICANT: LAFFER, SYLVIA

TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF,

```
;
; TITLE OF INVENTION: CORRESPONDING RECOMBINANT DNA MOLECULES, VECTORS AND HOSTS
; TITLE OF INVENTION: CONTAINING THE DNA MOLECULES, DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES OF SAID ALLERGENS AND FRAGMENTS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,419A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1614-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-750-419A-30

Query Match 22.3%; Score 40; DB 3; Length 29;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 RYVQKGVRTNPDIYP 22
   || :| :| ||:|
DB 1 RYTEGGTKTEADVIP 17

RESULT 5
US-08-750-419A-28
; Sequence 28, Application US/08750419A
; Patent No. 6008340
; GENERAL INFORMATION:
; APPLICANT: BALL, TANJA
; APPLICANT: VITALA, SUSANNE
; APPLICANT: SPERR, WOLFGANG
; APPLICANT: VALENT, PETER
; APPLICANT: SUSANT, MARKUS
; APPLICANT: KRAFT, DIETRICH
; APPLICANT: LAFFER, SYLVIA
; TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF,
; TITLE OF INVENTION: CORRESPONDING RECOMBINANT DNA MOLECULES, VECTORS AND HOSTS
; TITLE OF INVENTION: CONTAINING THE DNA MOLECULES, DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES OF SAID ALLERGENS AND FRAGMENTS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,419A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1614-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-750-419A-28

Query Match 22.3%; Score 40; DB 3; Length 30;
Best Local Similarity 41.2%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 RYVQKGVRTNPDIYP 22
   || :| :| ||:|
DB 2 RYTEGGTKTEADVIP 18

RESULT 6
US-09-215-221-11
; Sequence 11, Application US/09215221
; Patent No. 6265562
; GENERAL INFORMATION:
; APPLICANT: EILERS, MARTIN
; APPLICANT: BUERGIN, ANDREA
; APPLICANT: SEDLACEK, HANS-HARALD
; TITLE OF INVENTION: NUCLEIC ACID CONSTRUCTS WHOSE ACTIVITY IS AFFECTED BY
; TITLE OF INVENTION: INHIBITORS OF CYCLIN-DEPENDANT KINASES AND USES THEREOF
; FILE REFERENCE: 026083/0192
; CURRENT APPLICATION NUMBER: US/09/215,221
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 197 56 975.7
; PRIOR FILING DATE: 1997-12-20
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-215-221-11

Query Match 21.8%; Score 39; DB 4; Length 24;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 19 DIYPSNPTDDDV 30
   :|||:||||
DB 13 ETYDSNESDDDV 24

RESULT 7
US-08-446-692-8
; Sequence 8, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
```



;; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
;; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
;; NUMBER OF SEQUENCES: 114  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Maria C.H. Lin  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: US  
;; ZIP: 10154-0053  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/446,692  
;; FILING DATE: 7-JUN-1995  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maria C.H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)415-8745  
;; TELEFAX: (516)751-6849  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 24 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-446-692-8

Query Match 20.7%; Score 37; DB 1; Length 24;  
Best Local Similarity 44.4%; Pred. No. 39;  
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 CRDGRYVQKGEYRTNPE 18  
| :| | | : | | | |  
Db 6 CPNGTRALTVAELRGNAE 23

## RESULT 8

US-08-488-351A-8  
; Sequence 8, Application US/08488351A  
; Patent No. 5843446  
; GENERAL INFORMATION:

;; APPLICANT: Ladd, Anna  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
;; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
;; NUMBER OF SEQUENCES: 114  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Maria C.H. Lin  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: US  
;; ZIP: 10154-0053  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/488,351A  
;; FILING DATE: 7-JUN-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/446,692

;; FILING DATE: 7-JUN-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/229,275  
;; FILING DATE: 14-APR-1994  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/057,166  
;; FILING DATE: 27-APR-1992  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maria C.H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)415-8745  
;; TELEFAX: (516)751-6849  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 24 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-488-351A-8

Query Match 20.7%; Score 37; DB 2; Length 24;  
Best Local Similarity 44.4%; Pred. No. 39;  
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 CRDGRYVQKGEYRTNPE 18  
| :| | | : | | | |  
Db 6 CPNGTRALTVAELRGNAE 23

## RESULT 9

US-09-100-409A-44  
; Sequence 44, Application US/09100409A  
; Patent No. 6090388  
; GENERAL INFORMATION:

;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR  
;; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND  
;; TITLE OF INVENTION: IMMUNE DISORDERS  
;; NUMBER OF SEQUENCES: 64  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MORGAN & FINNEGAN  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154-0054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version  
;; SOFTWARE: #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/100,409A  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME:  
;; REGISTRATION NUMBER:  
;; REFERENCE/DOCKET NUMBER: 1151-4154  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-758-4800  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 44:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 24 amino acids  
;; TYPE: amino acid

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-44

Query Match 20.7%; Score 37; DB 3; Length 24;
Best Local Similarity 44.4%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CRDGTTRYVOKGYRTNPE 18
Db 6 CPNGTRALTVAELRGNAE 23

RESULT 10
PCT-US95-13841-11
; Sequence 11, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-11

Query Match 20.7%; Score 37; DB 5; Length 24;
Best Local Similarity 44.4%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CRDGTTRYVOKGYRTNPE 18
Db 6 CPNGTRALTVAELRGNAE 23

RESULT 11
US-08-563-892A-10
; Sequence 10, Application US/08563892A
; Patent No. 5976819

; GENERAL INFORMATION:
; APPLICANT: Finkel, Terri H.
; APPLICANT: Rozdzial, Moshe M.
; TITLE OF INVENTION: PRODUCT AND PROCESS TO REGULATE ACTIN
; POLYMERIZATION IN T LYMPHOCYTES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/563,892A
; FILING DATE: 21-NOV-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-563-892A-10

Query Match 20.7%; Score 37; DB 2; Length 32;
Best Local Similarity 47.4%; Pred. No. 55;
Matches 9; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

Qy 11 GEYRTNPEDIYPSNPTDDD 29
Db 11 GGYMT---LNPRAPTDDD 25

RESULT 12
US-08-938-595-5
; Sequence 5, Application US/08938595
; Patent No. 6157946
; GENERAL INFORMATION:
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Garner, Ian
; TITLE OF INVENTION: Peptide Production as Fusion Protein in
; Transgenic Mammal Milk
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,595
; FILING DATE: (Herewith)
```

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/727,153  
FILING DATE: 08-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/718,523  
FILING DATE: 08-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB95/00769  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9406974.7  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fleshner, Raz E.  
REGISTRATION NUMBER: 34,331  
REFERENCE/DOCKET NUMBER: 0623.0460003  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-938-595-5

Query Match 19.6%; Score 35; DB 4; Length 12;  
Best Local Similarity 75.0%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 22 PSNPTDDD 29  
I: |||||  
Db 3 PAFPTDDD 10

RESULT 13  
US-08-727-153-5  
Sequence 5, Application US/08727153  
Patent No. 6211427  
GENERAL INFORMATION:  
APPLICANT: Cottingham, Ian R.  
APPLICANT: Garner, Ian  
TITLE OF INVENTION: Peptide Production as Fusion Protein in  
TITLE OF INVENTION: Transgenic Mammal Milk  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/727,153  
FILING DATE: 08-OCT-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/718,523  
FILING DATE: 08-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB95/00769  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9406974.7

FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fleshner, Raz E.  
REGISTRATION NUMBER: 34,331  
REFERENCE/DOCKET NUMBER: 0623.0460001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-727-153-5

Query Match 19.6%; Score 35; DB 4; Length 12;  
Best Local Similarity 75.0%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 22 PSNPTDDD 29  
I: |||||  
Db 3 PAFPTDDD 10

RESULT 14  
US-08-440-861-52  
Sequence 52, Application US/08440861  
Patent No. 5710126  
GENERAL INFORMATION:  
APPLICANT: Griffith, Irwin J.  
APPLICANT: Kuo, Mei-Chang  
APPLICANT: Lugman, Mohammad  
TITLE OF INVENTION: T CELL EPITOPES OF RYEGRASS POLLEN  
TITLE OF INVENTION: ALLERGEN  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,861  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/106,016  
FILING DATE: 31-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-075 (IMI-040cp)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-440-861-52

Query Match 19.6%; Score 35; DB 1; Length 20;  
Best Local Similarity 35.3%; Pred. No. 62;  
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 RYVQKGEYRTNPDIYP 22  
||:|:|:|:|  
Db 2 RYVTEGTRKSEVEDVIP 18

## RESULT 15

US-07-729-099-15

; Sequence 15, Application US/07729099

; Patent No. 5403581

; GENERAL INFORMATION:

; APPLICANT: Binger, Mary-Helen

; APPLICANT: Pasamontes, Luis

; TITLE OF INVENTION: Coccidiosis Vaccines

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: New Jersey

; COUNTRY: U.S.

; ZIP: 07110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/729,099

; FILING DATE: 19910712

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Roseman, Catherine R

; REGISTRATION NUMBER: 34,240

; REFERENCE/DOCKET NUMBER: 8514

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-6208

; TELEFAX: (201) 235-3500

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 18 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: YES

; FRAGMENT TYPE: C-terminal

; ORIGINAL SOURCE:

; ORGANISM: Eimeria tenella

US-07-729-099-15

Query Match 19.0%; Score 34; DB 1; Length 18;  
Best Local Similarity 40.0%; Pred. No. 77;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 15 TNPDYPSNPTDDD 29  
||:|:|:|:|  
Db 1 TQPDTEFRSGPGDDE 15

Search completed: February 21, 2002, 16:34:47  
Job time: 112 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:37:20 ; Search time 13.26 seconds  
(without alignments)  
241.277 Million cell updates/sec

Title: US-08-753-851-4  
Perfect score: 223  
Sequence: 1 CRDGRVQKGEYRINPEDI.....SNPTDDVSSGSSRSSTS 42

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 9284

Minimum DB seq length: 0  
Maximum DB seq length: 42

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	33.6	29	2	A41683
2	40	17.9	29	2	I49732
3	40	17.9	30	2	PC2307
4	37	16.6	38	2	S40096
5	30.5	16.4	38	2	A44862
6	36	16.1	35	2	S54329
7	36	16.1	40	2	T07488
8	35.5	15.9	31	2	C39513
9	33	14.8	19	2	A41299
10	33	14.8	26	2	G37396
11	33	14.8	28	2	A03356
12	33	14.8	34	2	E45357
13	33	14.8	40	2	A29184
14	32	14.3	20	2	S77983
15	32	14.3	24	2	C45357
16	32	14.3	28	2	PS0106
17	32	14.3	29	2	S35924
18	32	14.3	36	2	S35572
19	32	14.3	38	2	A32112
20	32	14.3	38	2	S78757
21	32	14.3	40	2	PQ0050
22	31	13.9	40	2	B27398
23	30.5	13.7	36	2	A40723
24	30.5	13.7	36	2	F81745
25	30.5	13.7	42	2	S78722
26	30	13.5	19	2	S28396
27	30	13.5	20	2	G61491
28	30	13.5	22	2	I50533
29	30	13.5	30	2	A84412

30 30 13.5 35 2 S07052 neuroptide Y - s  
31 30 13.5 39 1 CTDFAS corticotropin - sp  
32 29.5 13.2 19 2 A61377 endometrial secret  
33 29.5 13.2 20 2 S32387 ribosomal protein  
34 29.5 13.2 40 2 A05301 hemoglobin alpha-2  
35 29 13.0 18 2 A32917 protein phosphatas  
36 29 13.0 25 2 PQ0369 L protein - rabies  
37 29 13.0 25 2 T01689 ATPase-beta chain  
38 29 13.0 26 2 F37396 pollen allergen Po  
39 29 13.0 26 2 E37396 pollen allergen Ag  
40 29 13.0 26 2 D53507 annexin V - rat (f  
41 29 13.0 29 2 B56817 photosystem I chai  
42 29 13.0 30 2 S13994 hypothetical prote  
43 29 13.0 31 2 T01701 hypothetical prote  
44 29 13.0 34 2 F82394 hypothetical prote  
45 29 13.0 34 2 B45357 kex2/subtilisin-li

## ALIGNMENTS

RESULT 1  
A41683  
hyaluronate receptor CD44 (version 2) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: A41683  
R:Shtivelman, E.; Bishop, J.M.  
Mol. Cell. Biol. 11, 5446-5453, 1991  
A:Title: Expression of CD44 is repressed in neuroblastoma cells.  
A:Reference number: A41683; MUID:92017823  
A:Accession: A41683  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-29 <SHT>  
A:Cross-references: GB:M69215  
C:Keywords: cell adhesion

Query Match 33.6% Score 75 DB 2: Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 DDDVSSGSSRSSTS 42  
DB 1 DDDVSSGSSRSSTS 16  
|||||

## RESULT 2

I49732  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) 14K chain - potato mitochondrion (fragme  
N:Alternate names: complex 1 dehydrogenase 14K chain; NADH-ubiquinone oxidoreductase  
C:Species: mitochondrion Solanum tuberosum (potato)  
C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 07-May-1999  
C:Accession: I49732  
R:Herz, U.; Schroeder, W.; Liddell, A.; Leaver, C.J.; Brennicke, A.; Grohmann, L.  
J. Biol. Chem. 269, 2263-2269, 1994  
A:Title: Purification of the NADH:ubiquinone oxidoreductase (complex I) of the respir  
A:Reference number: A49732; MUID:94124587  
A:Accession: I49732  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-29 <HER>  
C:Genetics:  
A:Genome: mitochondrion  
C:Keywords: electron transfer; hydrogen ion transport; mitochondrial inner membrane;

Query Match 17.9% Score 40 DB 2: Length 29;  
Best Local Similarity 46.7%; Pred. No. 89;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 17 PEDIYPSNPTDDVS 31



RESULT 10  
G37396  
pollen allergen Ant o I - sweet vernal grass (fragment)  
C:Species: Anthoxanthum odoratum (sweet vernal grass)  
C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 09-Jun-2000  
C:Accession: G37396  
R:Esch, R.E.; Klapper, D.G.  
Mol. Immunol. 26, 557-561, 1989  
A:Title: Isolation and characterization of a major cross-reactive grass group I allergen  
A:Reference number: A37396; MUID:89364850  
A:Accession: G37396  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-26 <ESC>  
C:Superfamily: expansin  
C:Keywords: pollen

Job time: 100 sec

A:Molecule type: protein  
A:Residues: 1-40 <CLA>  
C:Superfamily: vitellogenin  
C:Keywords: phosphoprotein  
F:12,13,14,15,16,17,19,20,24,25,26,27,28,29,31,32/Binding site: phosphate (Ser) (covalent)

Query Match 14.8%; Score 33; DB 2; Length 40;  
Best Local Similarity 42.9%; Pred. No. 1e+03;  
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 22 PSNPTDDVSSGSSRSSTS 42  
| | | | | | | | | |  
DB 7 PNKTSSTSSSVSTATSSS 27

## RESULT 14

S77983  
cytochrome-c oxidase (EC 1.9.3.1) chain Vb - bigeye tuna (fragment)  
C:Species: Thunnus obesus (bigeye tuna)  
C:Date: 17-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 30-Jan-1998  
C:Accession: S77983  
R:Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.  
submitted to the Protein Sequence Database, June 1997  
A:Reference number: S77980  
A:Accession: S77983  
A:Molecule type: protein  
A:Residues: 1-20 <ARN>  
A:Experimental source: heart; liver  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Pathway: oxidative phosphorylation; respiratory chain  
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 14.3%; Score 32; DB 2; Length 20;  
Best Local Similarity 55.6%; Pred. No. 6.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 25 PTDDDVSSG 33  
| | | | | | | | | |  
DB 7 PTDEQATG 15

## RESULT 15

C45357  
kex2/subtilisin-like proprotein convertase PC4-C - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
C:Accession: C45357  
R:Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.  
Mol. Endocrinol. 6, 1359-1370, 1992  
A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell protein  
A:Reference number: A45357; MUID:93078790  
A:Accession: C45357  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-24 <SEI>  
A>Note: sequence extracted from NCBI backbone (NCBIP:118879)

Query Match 14.3%; Score 32; DB 2; Length 24;  
Best Local Similarity 33.3%; Pred. No. 7.6e+02;  
Matches 7; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 22 PSNPTDDVSSGSSRSSTS 42  
| | | | | | | | | |  
DB 3 PQNSASSPASSGGTATHSS 23

Search completed: February 21, 2002, 16:39:00





Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	40	17.9	29	1	NUOL_SOLTU	P80267	solanum tub
2	33.5	15.0	37	1	YL28_CAEEL	P34421	caenorhabdi
3	33	14.8	28	1	GDO_TRIMO	P02865	triticum mon
4	33	14.8	40	1	VIT_MEIGA	P56531	meleagris g
5	32	14.3	20	1	COXB_THUOB	P80974	thunnius obe
6	31	13.9	36	1	PAHO_RABIT	P41336	oryctolagus
7	31	13.9	40	1	PHAC_MASLA	P11389	mastigoclad
8	30	13.5	36	1	NEUY_SHEEP	P14765	ovis aries
9	30	13.5	36	1	PYY_RAURH	P29206	raja rhina
10	30	13.5	39	1	COLI_SQUAC	P01197	squalus aca
11	29.5	13.2	40	1	HBA2_UROHA	P18980	uromastyx h
12	29	13.0	28	1	PSAF_SNP6	P31083	synechococ
13	29	13.0	39	1	VG8_SPV4	P11340	spiroplasma
14	29	13.0	39	1	COLI_BALPH	P01195	baiaenopter
15	28.5	12.8	37	1	REV_SIVM2	P08809	simian immu
16	28.5	12.8	37	1	COLI_STRCA	P01196	struthio car
17	28	12.6	17	1	LCK_RAT	Q01621	rattus norv
18	28	12.6	20	1	COXF_ONCMY	P80329	oncorhynch
19	28	12.6	24	1	FEBG_ANYME	P80707	amycotatops
20	28	12.6	28	1	PHAC_ANASP	P80556	anabaena sp
21	28	12.6	29	1	HOXY_RHOOP	P22660	rhodococcus
22	28	12.6	30	1	CY35_DESAC	P81079	desulfurosom
23	28	12.6	33	1	COXL_ONCMY	P80330	oncorhynch
24	28	12.6	36	1	NEUY_GADMO	P80167	gadus morhua
25	28	12.6	36	1	NEUY_ONCMY	P29071	oncorhynch
26	28	12.6	39	1	NPF_MONEX	P41967	monelia ex
27	27.5	12.3	16	1	MK2B_PALPR	P80410	palomena pr
28	27.5	12.3	33	1	SCX9_BUTOC	P04099	buthus occi
29	27.5	12.3	34	1	DEF7_RABIT	P80223	oryctolagus
30	27.5	12.3	40	1	YDRB_STRPE	P32012	streptomyces
31	27	12.1	18	1	YMOB_CHLAI	Q45827	chloroflexu
32	27	12.1	23	1	ECDB_LYMDI	P80939	lymantria d
33	27	12.1	25	1	COX6_NEUCR	Q01359	neurospora

RESULT	2			
YLZ8_CAEEL				
ID	YLZ8_CAEEL	STANDARD:	PRT;	37 AA.
AC	P34421;			
DT	01-FEB-1994	(Rel. 28, Created)		
DT	01-FEB-1994	(Rel. 28, Last sequence update)		
DT	01-JUN-1994	(Rel. 28, Last annotation update)		
DE	HYPOTHETICAL 4.4 KDA PROTEIN F42H10.8 IN CHROMOSOME III.			
GN	F42H10.8.			
OS	Caenorhabditis elegans.			

```
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Stulton J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08403; ; NOT_ANNOTATED_CDS.
DR WormPep; F42H10.8; CE00358.
KW Hypothetical protein
SQ SEQUENCE 37 AA; 4433 MW; 1F7C5F19F020105C CRC64;

Query Match
Best Local Similarity 15.0%; Score 33.5; DB 1; Length 37;
Matches 8; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 13 YRNPEDIVPSNPTDDVS 31
Db :| | | | |
18 HKILP-DLYPLEPQQAVS 35

RESULT 3
GDO_TRIMO
ID GDO_TRIMO STANDARD; PRT; 28 AA.
AC P02865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE OMEGA-GLIADIN (FRAGMENT).
OS Triticum monococcum (Einkorn wheat) (Small spelt).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4588;
RN [1]
RP SEQUENCE.
RA Shewry P.R., Autran J.-C., Nimmo C.C., Lew E.J.-L., Kasarda D.D.;
RT "N-terminal amino acid sequence homology of storage protein components
RT from barley and a diploid wheat."
RL Nature 286:520-522(1980).
DR PIR; A03356; A03356.
KW Seed storage protein.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3343 MW; EA368C84E893FAA7 CRC64;

Query Match
Best Local Similarity 14.8%; Score 33; DB 1; Length 28;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 16 NPEDIYSPNP 25
Db :| | | |
13 SPQQLYPOQP 22

RESULT 4
VIT_MELGA
ID VIT_MELGA STANDARD; PRT; 40 AA.
AC P56531;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PHOSVITIN (PV) (FRAGMENT).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE.
RX MEDLINE=86056531; PubMed=4065410;
RA Clark R.C.;
RT "The primary structure of avian phosvitins. Contributions through the
RT Edman degradation of methylmercaptovitins prepared from the
RT constituent phosphoproteins."
RL Int. J. Biochem. 17:983-988(1985).
CC -1- FUNCTION: PHOSVITIN IS BELIEVED TO BE OF IMPORTANCE IN
CC SEQUESTERING CALCIUM, IRON AND OTHER CATIONS FOR THE DEVELOPING
CC EMBRYO.
CC -1- PTM: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST
CC HIGHLY PHOSPHORYLATED (10%) PROTEINS IN NATURE.
KW Phosphorylation; Storage protein.
FT NON_TER 1
FT DOMAIN 12 20 POLY-SER.
FT DOMAIN 24 32 POLY-SER.
FT NON_TER 40 40
SQ SEQUENCE 40 AA; 3935 MW; F736301BDCFE708B CRC64;

Query Match
Best Local Similarity 14.8%; Score 33; DB 1; Length 40;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 22 PSNPTDDVSSGSSSRSSTS 42
Db :| | | | |
7 PNKTSSTSSSVSSTATSSSS 27

RESULT 5
COXB_THUOB
ID COXB_THUOB STANDARD; PRT; 20 AA.
AC P80974;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VB (EC 1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver."
RL Eur. J. Biochem. 248:99-103(1997).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
```

```

CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.
DR InterPro: IPR002124; COX5B.
DR PROSITE: PS00848; COX5B; PARTIAL.
KW Oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2158 MW; 76F1473ELF392BD7 CRC64;

Query Match 14.3%; Score 32; DB 1; Length 20;
Best Local Similarity 55.6%; Pred. No. 2.8e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 25 PTDDDVSSG 33
| | | | |
DB 7 PTDEQATG 15

RESULT 6
ID PAHO_RABIT STANDARD; PRT; 36 AA.
AC P41336.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PANCREATIC HORMONE (PANCREATIC POLYPEPTIDE) (PP).
GN PPY.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=94130533; PubMed=8299350;
RA Marks N.J., Shaw C., Halton D.W., Curry W.J., Thim L.;
RT "Rabbit pancreatic polypeptide."
RL Comp. Biochem. Physiol. 106B:883-887(1993).
CC -!- FUNCTION: PANCREATIC HORMONE IS SYNTHESIZED IN PANCREATIC ISLETS
CC OF LANGERHANS AND ACTS AS A REGULATOR OF PANCREATIC AND
CC GASTROINTESTINAL FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.
DR HSSP; P01302; 1BBA.
DR InterPro: IPR001955; Pancreatic_hormn.
DR Pfam; PF00159; hormone3; 1.
DR PRINTS; PR00278; PANCHORMONE.
DR ProDom; PD001267; Pancreatic_hormn; 1.
DR SMART; SM00309; PAH; 1.
DR PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
DR PROSITE; PS0276; PANCREATIC_HORMONE_2; 1.
KW Hormone; Amidation; Pancreas.
FT MOD_RES 36 36
FT SEQUENCE 36 AA; 4197 MW; A14A4E0831A7759D CRC64;

Query Match 13.9%; Score 31; DB 1; Length 36;
Best Local Similarity 33.3%; Pred. No. 7.5e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 4; Gaps 1;

QY 17 PEDTYPSTDDVSSGSSSE 37
| | | | |
DB 3 PEFVYPG----DDATPEQMAE 19

RESULT 7
ID PHAC_MASLA STANDARD; PRT; 40 AA.
AC P11389.
DT 01-JUL-1989 (Rel. 11, Created)

```

```

DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE ALLOPHYCOCYANIN ALPHA-B CHAIN (FRAGMENT).
OS Mastigocladus laminosus (Fischerella sp.).
OC Bacteria; Cyanobacteria; Stigonematales; Fischerella.
OX NCBI_TaxID=1191;
RN [1]
RP SEQUENCE.
RA Suter F., Fueglistaller P., Lundell D.J., Glazer A.N., Zuber H.;
RT "Amino acid sequences of alpha-allophycocyanin B from Synechococcus
RT 6301 and Mastigocladus laminosus."
RL FEBS Lett. 217:279-282(1987).
CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
DR InterPro: IPR001659; Phycobilisome.
DR PIR; B27398; B27398.
DR Pfam; PF00502; Phycobilisome; 1.
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment.
FT NON_TER 40 40
FT SEQUENCE 40 AA; 4506 MW; 2FC4F66D3F091D54 CRC64;

Query Match 13.9%; Score 31; DB 1; Length 40;
Best Local Similarity 45.5%; Pred. No. 8.4e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 TRYVQGEYRI 15
| | | | |
DB 26 TEFLTQGEQRV 36

RESULT 8
ID NEUY_SHEEP STANDARD; PRT; 36 AA.
AC P14765;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROPEPTIDE Y (NPY).
GN NPY.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=90092485; PubMed=2599092;
RA Sillard R., Agerberth B., Mutt V., Joernvall H.;
RT "Sheep neuropeptide Y. A third structural type of a highly conserved
RT peptide."
RL FEBS Lett. 258:263-265(1989).
CC -!- FUNCTION: NPY IS IMPLICATED IN THE CONTROL OF FEEDING AND IN
CC SECRETION OF GONADOTROPHIN-RELEASE HORMONE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: ONE OF THE MOST ABUNDANT PEPTIDES IN THE
CC NERVOUS SYSTEM. ALSO FOUND IN SOME CHROMAFFIN CELLS OF THE ADRENAL
CC MEDULLA.
CC -!- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.
DR PIR; S07052; S07052.
DR HSSP; P01303; IRON.
DR InterPro: IPR001955; Pancreatic_hormn.
DR Pfam; PF00159; hormone3; 1.
DR PRINTS; PR00278; PANCHORMONE.
DR ProDom; PD001267; Pancreatic_hormn; 1.
DR SMART; SM00309; PAH; 1.
DR PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
DR PROSITE; PS0276; PANCREATIC_HORMONE_2; 1.
KW Neuropeptide; Amidation.
FT MOD_RES 36 36
FT SEQUENCE 36 36

```



Query Match 13.2%; Score 29.5; DB 1; Length 40;  
Best Local Similarity 35.0%; Pred. No. 1.3e+03;  
Matches 7; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 19 DIYPSNPT---DDVSSGSS 35  
DB 18 EVYPTQKTYFPHPDASAGSN 37

RESULT 12  
PSAF\_SYN6 STANDARD; PRT; 29 AA.  
AC P31083;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT III (PSI-F) (FRAGMENT).  
GN PSF.  
OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
OX NCBI\_TaxID=1139;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=91355213; PubMed=1653017;  
RA Li N., Warren P.V., Golbeck J.H., Frank G., Zuber H., Bryant D.A.;  
RT "Polypeptide composition of the Photosystem I complex and the  
RT Photosystem I core protein from Synechococcus sp. PCC 6301.";  
RL Biochim. Biophys. Acta 1059:215-225(1991).  
CC -!- FUNCTION: PROBABLY PARTICIPATES IN EFFICIENCY OF ELECTRON TRANSFER  
CC FROM PLASTOCYANIN TO P700 (OR CYTOCHROME C553 IN ALGAE AND  
CC CYANOBACTERIA). THIS PLASTOCYANIN-DOCKING PROTEIN CONTRIBUTES  
CC TO THE SPECIFIC ASSOCIATION OF PLASTOCYANIN TO PS I.  
CC -!- SIMILARITY: BELONGS TO THE PSF FAMILY.  
KW Photosynthesis; Photosystem I.  
FT NON\_TER 29  
SQ SEQUENCE 29 AA; 3040 MW; 535DDC7A63B18711 CRC64;

Query Match 13.0%; Score 29; DB 1; Length 29;  
Best Local Similarity 25.0%; Pred. No. 1e+03;  
Matches 4; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 DGTYYVKGXEYRINPE 18  
DB 10 ESPRFQRAEAAATPQ 25

RESULT 13  
VG8\_SPV4 STANDARD; PRT; 38 AA.  
AC P11340;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 01-JUL-1989 (Rel. 11, Last annotation update)  
DE GENE 8 PROTEIN.  
GN 8.

OS Spiroplasma virus 4 (Spv4).  
OC Viruses; ssDNA viruses; Microviridae; Spiromicrovirus.  
OX NCBI\_TaxID=10855;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88032809; PubMed=2822658;  
RA Renaudin J., Pascarel M.-C., Bove J.-M.;  
RT "Spiroplasma virus 4: nucleotide sequence of the viral DNA,  
RT regulatory signals, and proposed genome organization.";  
RL J. Bacteriol. 169:4950-4961(1987).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; M17988; -, NOT\_ANNOTATED\_CDS.  
DR PIR; A29825; G8PSV.  
SQ SEQUENCE 38 AA; 4629 MW; 58E00837B14DF2C2 CRC64;

Query Match 13.0%; Score 29; DB 1; Length 38;  
Best Local Similarity 54.5%; Pred. No. 1.4e+03;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 INPEIYPSNP 25  
DB 22 IKRANIMPSNP 32

RESULT 14  
COLI\_BALPH STANDARD; PRT; 39 AA.  
ID COLI\_BALPH  
AC P01195;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CORTICOTROPIN-LIPOTROPIN (PRO-OPIOMELANOCORTIN) (POMC) [CONTAINS:  
DE CORTICOTROPIN (ADRENOCORTICOTROPIC HORMONE) (ACTH); MELANOTROPIN ALPHA  
DE (ALPHA-MSH); CORTICOTROPIN-LIKE INTERMEDIARY PEPTIDE (CLIP)]  
DE (FRAGMENT).  
GN POMC.  
OS Balaenoptera physalus (Finback whale) (Common rorqual), and  
OS Balaenoptera borealis (Sei whale).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
OC Balaenopteridae; Balaenoptera.  
OX NCBI\_TaxID=9770, 9768;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=B. physalus;  
RA Pankov Y.A., Nikolaeva O.P., Elisarova G.P.;  
RT "Primary structure of the corticotropin of whalebone whales, finbacks  
RT (Balaenoptera physalus).";  
RL Bioorg. Khim. 2:855-856(1976).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=B. borealis;  
RX MEDLINE=78061046; PubMed=201308;  
RA Pankov Y.A., Nikolaeva O.P., Elisarova G.P.;  
RT "Amino acid sequence of corticotropins from seiwhale (Balaenoptera  
RT borealis) and pinwhale (Balaenoptera physalus).";  
RL Biochimica 42:2044-2050(1977).  
CC -!- FUNCTION: ACTH STIMULATES THE ADRENAL GLANDS TO RELEASE CORTISOL.  
CC -!- FUNCTION: MSH (MELANOCYTE-STIMULATING HORMONE) INCREASES THE  
CC PIGMENTATION OF SKIN BY INCREASING MELANIN PRODUCTION IN  
CC MELANOCYTES.  
CC -!- FUNCTION: BETA-ENDORPHIN AND MET-ENKEPHALIN ARE ENDOGENOUS  
CC OPIATES.  
CC -!- TISSUE SPECIFICITY: ACTH AND MSH ARE PRODUCED BY THE PITUITARY  
CC GLAND.  
CC -!- SIMILARITY: BELONGS TO THE POMC FAMILY.  
DR PIR; A01458; A01458.  
DR PIR; P0127; P0127.  
DR InterPro; IPR001941; Melanocortin\_ACTH.  
DR Pfam; PF00976; ACTH\_domain; 1.  
DR ProDom; PD003250; Melanocortin\_ACTH; 1.  
KW Hormone; Cleavage on pair of basic residues; Amidation.  
FT NON\_TER 1 1  
FT PEPTIDE 1 39 CORTICOTROPIN.  
FT PEPTIDE 1 13 MELANOTROPIN ALPHA.  
FT PEPTIDE 19 39 CORTICOTROPIN-LIKE INTERMEDIARY PEPTIDE.  
FT MOD\_RES 13 13 AMIDATION (G-14 PROVIDE AMIDE GROUP).  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4541 MW; 461FD46C28945A6E CRC64;

Query Match 13.0%; Score 29; DB 1; Length 39;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 17 PEDIYSPNPTDDVSS 32  
 | :||: |:  
 Db 19 PVKVPNGAEDESAEA 34

## RESULT 15

REV\_SIVM2  
 ID REV\_SIVM2 STANDARD; PRT; 37 AA.  
 AC P08809;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DE REV PROTEIN (ANTI-REPRESSION TRANSACTIVATOR PROTEIN) (ART/TRS)  
 DE (FRAGMENT).  
 GN REV.  
 OS Simian immunodeficiency virus (Mm251 isolate) (SIV-MAC).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11734;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88122665; PubMed=2893293;  
 RA Kestler H.W., Li Y., Naidu Y.M., Butler C.V., Ochs M.F., Jaenel G.,  
 RA King N.W., Daniel M.D., Destroiers R.C.;  
 RT "Comparison of simian immunodeficiency virus isolates.";  
 RL Nature 331:619-622(1988).  
 CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE  
 CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR; ACCUMULATES IN THE NUCLEOLI.  
 CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED  
 CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.  
 CC -!- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X06879; -; NOT\_ANNOTATED\_CDS.  
 DR HIV; X06879; REV\$MM251.  
 DR InterPro; IPR000625; REV.  
 DR Pfam; PF00424; REV; 1.  
 KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 37 37  
 SQ SEQUENCE 37 AA; 4633 MW; 059C315CC56C5583 CRC64;

Query Match 12.8%; Score 28.5; DB 1; Length 37;  
 Best Local Similarity 30.0%; Pred. No. 1.6e+03;  
 Matches 6; Conservative 5; Mismatches 2; Indels 7; Gaps 1;

QY 19 DIYPSNPTDDVSSGSSSR 38  
 | ||: |  
 Db 1 DSYPTGP-----GTANQR 13

Search completed: February 21, 2002, 16:41:45  
 Job time: 179 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:38:25 ; Search time 21.83 Seconds

(without alignments)  
281.422 Million cell updates/sec

Title: US-08-753-851-4

Perfect score: 223

Sequence: 1 CRGTRVQKGEYRINPEDI.....SNPTDDVSSGSSRSSTS 42

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 21122

Minimum DB seq length: 0

Maximum DB seq length: 42

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.5	21.7	38	5	O96850 bombyx mori
2	40	17.9	30	11	Q9QUV6 cavia (guin
3	37	16.6	39	10	O40959 physcomitre
4	36.5	16.4	38	5	Q26899 trypanosoma
5	36.5	16.4	38	5	Q26343 trypanosoma
6	36	16.1	40	8	Q32954 pinus thunb
7	35.5	15.9	42	4	Q9BTP5 homo sapien
8	35	15.7	20	5	Q9TWI5 ceratitidis c
9	35	15.7	42	5	Q9VSI7 drosophila
10	34.5	15.5	38	5	O96852 bombyx mori
11	34	15.2	20	5	Q9TWI8 ceratitidis c
12	34	15.2	28	12	Q9EIT3 rabies viru
13	34	15.2	42	3	Q9P7A7 schizosacch
14	34	15.2	42	4	Q9UH60 homo sapien
15	33.5	15.0	38	2	Q9R5U8 agrobacteri
16	33	14.8	22	1	Q9UWK5 methanobact
17	33	14.8	26	2	O86139 clostridium
18	33	14.8	26	5	Q9U6U9 plasmodium
19	33	14.8	28	12	Q9EIT4 rabies viru

Q9UCS9 homo sapien  
Q9WRL4 avian pneum  
Q9KZ25 streptomyce  
Q9TWE3 ascaris suu  
Q9QH0 rattus norv  
Q9CRF6 mus musculu  
Q91866 porcine cir  
Q96851 bombyx mori  
Q95371 paramecium  
Q93618 streptomyce  
Q93363 chlamydia t  
Q9UEE7 homo sapien  
Q9NQ77 homo sapien  
Q93794 salmonella  
Q9NLC6 ovis aries  
Q91885 photobacter  
Q9NR12 homo sapien  
Q9TX41 cryptospori  
Q9QW82 rattus norv  
O18764 sus scrofa  
Q9TNS5 homo sapien  
Q9QD09 rabies viru  
Q9R4J2 helicobacte  
Q86818 influenza a  
Q9KI18 anabaena sp  
Q99138 avian influ

20 33 14.8 35 4 Q9UCS9  
21 33 14.8 35 12 Q9WRL4  
22 33 14.8 40 2 Q9KZ25  
23 33 14.8 42 5 Q9TWE3  
24 33 14.8 42 11 Q9QH0  
25 32.5 14.6 27 11 Q9CRF6  
26 32.5 14.6 29 12 Q91866  
27 32.5 14.6 38 5 Q96851  
28 32.5 14.6 40 8 Q95371  
29 32 14.3 30 2 Q93618  
30 32 14.3 31 2 Q93363  
31 32 14.3 34 4 Q9UEE7  
32 32 14.3 35 4 Q9NQ77  
33 32 14.3 37 2 Q93794  
34 32 14.3 37 6 Q9NLC6  
35 32 14.3 40 2 Q91885  
36 32 14.3 40 4 Q9NR12  
37 31.5 14.1 41 5 Q9TX41  
38 31 13.9 18 11 Q9QW82  
39 31 13.9 21 6 O18764  
40 31 13.9 24 7 Q9TNS5  
41 31 13.9 28 12 Q9QD09  
42 31 13.9 30 2 Q9R4J2  
43 31 13.9 30 12 Q86818  
44 31 13.9 31 2 Q9KI18  
45 31 13.9 33 12 Q99138

## ALIGNMENTS

RESULT 1

O96850 PRELIMINARY; PRT; 38 AA.  
AC O96850;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE SERICIN MRNA, PARTIAL CDS (FRAGMENT).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Bombycoidea; Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87076763; PubMed=3024742;  
RA Michaille J.-J., Coublie P., Prudhomme J.-C., Garel A.;  
RT "A single gene produces multiple sericin messenger RNAs in the silk  
gland of Bombyx mori.";  
RL Biochimie 68:1165-1173(1986).  
DR EMBL; M26103; AAA27845.1; -;  
FT NON\_TER 1  
FT NON\_TER 38  
SQ SEQUENCE 38 AA; 3684 MW; 84FACB37E32C7A21 CRC64;

Query Match 21.7%; Score 48.5; DB 5; Length 38;  
Best Local Similarity 56.5%; Pred. No. 24;  
Matches 13; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 21 YPSNPTDDVSS-GSSRSRSSTS 42

| | | | | | | | | | | | | | | | | | | | | |

Db 13 YSSNRDSVSSGSSNTDSNS 35

RESULT 2

O9QUV6 PRELIMINARY; PRT; 30 AA.  
ID O9QUV6  
AC O9QUV6;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)



```

DE AMINOPEPTIDASE P (FRAGMENT).
OC Cavia (guinea pigs).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.
OX NCBI_TaxID=10140;
RN [1]
RP SEQUENCE.
RX MEDLINE=95110325; PubMed=7811266;
RA Denslow N.D., Ryan J.W., Nguyen H.P.;
RT "Guinea pig membrane-bound aminopeptidase P is a member of the proline
RT peptidase family.";
RL Biochem. Biophys. Res. Commun. 205:1790-1795(1994).
SQ SEQUENCE 30 AA; 3411 MW; A9F6AG1854E44A25 CRC64;

Query Match 17.9%; Score 40; DB 11; Length 30;
Best Local Similarity 46.7%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 YVQKGEYRINPEDIY 21
   | | | | | | | | | |
Db 10 YYQDGEIGIEDVF 24

RESULT 3
ID Q40959 PRELIMINARY; PRT; 39 AA.
AC Q40959;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CHLOROPHYLL-A-B-BINDING PROTEIN (FRAGMENT).
GN LHCO OR CAB.
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. (HEDW.)B.S.G.;
RX MEDLINE=94359457; PubMed=8078460;
RA Reski R., Faust M., Wang X.H., Wehe M., Abel W.O.;
RT "Genome analysis of the moss Physcomitrella patens (Hedw.) B.S.G.";
RL Mol. Gen. Genet. 244:352-359(1994).
DR EMBL: X76632; CAA54081.1;
DR Mendel; 182; Phyta; Lnc0; 182.
DR InterPro; IPR001344; Chloro_AB_bind.
DR Pfam; PF00504; chloroa_b-bind; 1.
KW Chlorophyll.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 3947 MW; 69BA0C1296FBC1EA CRC64;

Query Match 16.6%; Score 37; DB 10; Length 39;
Best Local Similarity 50.0%; Pred. No. 7.3e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 26 TDDVSSGSSSSRSST 41
   | | | | | | | | | |
Db 14 SDPETSPTGTSWRSSST 29

RESULT 4
ID Q26899 PRELIMINARY; PRT; 38 AA.
AC Q26899;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ANTIGEN DNA (TANDEM REPEAT SEQUENCE), CLONE 36 (FRAGMENT).
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

```

```

OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88288332; PubMed=3135494;
RA Ibanez C.F., Affranchino J.L., Macina R.A., Reyes M.B., Leguizamon S.,
RA Camargo M.E., Aslund L., Pettersson U., Frasch A.C.;
RT "Multiple Trypanosoma cruzi antigens containing tandemly repeated
RT amino acid sequence motifs.";
RL Mol. Biochem. Parasitol. 30:27-33(1988).
DR EMBL: M21331; AAA75430.1;
KW Repeat.
FT NON_TER 1
FT NON_TER 38
SQ SEQUENCE 38 AA; 4409 MW; 25DEA4DBCE11E8DA CRC64;

Query Match 16.4%; Score 36.5; DB 5; Length 38;
Best Local Similarity 33.3%; Pred. No. 8.2e+02;
Matches 8; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 4 GTRVYQKGEYRINPEDIY-PSNPT 26
   | | | | | | | | | |
Db 11 GPRHVDPDHFRSTQDAYRPVDP 34

RESULT 5
ID Q26343 PRELIMINARY; PRT; 38 AA.
AC Q26343;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE MICROTUBULE ASSOCIATED PROTEIN HOMOLOG (FRAGMENT).
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TULAHUEN 2;
RX MEDLINE=92070444; PubMed=1959572;
RA Kerner N., Liegeard P., Levin M.J., Hontebeyrie-Joskowicz M.;
RT "Trypanosoma cruzi: antibodies to a MAP-like protein in chronic
RT Chagas' disease cross-react with mammalian cytoskeleton.";
RL Exp. Parasitol. 73:451-459(1991).
DR EMBL: S68286; AAB20531.2;
FT NON_TER 38
SQ SEQUENCE 38 AA; 4428 MW; D53D1F64DBCDF19F CRC64;

Query Match 16.4%; Score 36.5; DB 5; Length 38;
Best Local Similarity 33.3%; Pred. No. 8.2e+02;
Matches 8; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 4 GTRVYQKGEYRINPEDIY-PSNPT 26
   | | | | | | | | | |
Db 10 GPRHVDPDHFRSTQDAYRPVDP 33

RESULT 6
ID Q32954 PRELIMINARY; PRT; 40 AA.
AC Q32954;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF40B.
GN PSAM.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3350;
RN [1]

```

DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT	01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE	MAJOR CUTICULAR PROTEIN (FRAGMENT).
OS	Ceratitis capitata (Mediterranean fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Prterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Tephritoides; Tephritidae; Ceratitidis.
OX	NCBI_TaxID=7213;
RN	[1]
RN	SEQUENCE.
RX	MEDLINE=952791142; PubMed=7759285;
RA	"Silliras A., Korantzou E., Allen G., Dimitriadis G.;
RT	"Partial N-terminal sequences of larval cuticular proteins from the
RL	dipteran <i>Ceratitis capitata</i> .";
RL	Hereditas 122:79-83(1995).
SQ	SEQUENCE 20 AA; 2329 MW; C6ABB0B44ED70B88 CRC64;

  

Query Match	15.7%;	Score 35;	DB 5;	Length 20;
Best Local Similarity	42.9%;	Pred. No. 6.4e+02;		
Matches	6;	Conservative	3;	Mismatches 5; Indels 0; Gaps

  

Qy	8 VQGEYRNPEDIY 21
Dd	I :   :    :
	6 VIRSESEVNPDSIY 19

  

RESULT	9
Q9VS17	
ID	Q9VS17 PRELIMINARY; PRT; 42 AA.
AC	Q9VS17;
DC	01-MAY-2000 (TReMBLrel. 13, Created)
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT	01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE	CG13666 PROTEIN.
GN	CG13666.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Prterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-BERKELEY;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berham B.P., Bhattacharya D., Bolshakov S.,
RA	Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,
RA	Burtis K.C., Busam D.R., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Furber K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ithegawam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclet J.M.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

RESULT 11
Q9TWI8
ID Q9TWI8 PRELIMINARY; PRT; 20 AA.
AC Q9TWI8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE MAJOR CUTICULAR PROTEIN (FRAGMENT).
OS Ceratitis capitata (Mediterranean fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Tephritoidea; Tephritidae; Ceratitidae.
NCBI_TaxID=7213;
RN [1];

```

RESULT	13
Q9P7A7	
ID	PRELIMINARY; PRS 42 AA.
AC	Q9P7A7:
DT	01-OCT-2000 (TREMblrel. 15, Created)
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT	01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE	HYPOTHETICAL SERINE-RICH PROTEIN (FRAGMENT).
GN	SPAC1705.03C.
OS	Schizosaccharomyces pombe (Fission yeast).
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;
OC	Schizosaccharomycetes.
NCBI_TaxID	4896;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=972H-;
RA	Brown S., Harris D., Wood V., Rajadream M.A., Barrell B.G.;
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; ALI3529; CAB86946.1; -
FT	NON_TER 42
SO	SEQUENCE 42 AA; 3953 MW; 262E24C5F241951C CRG64;

Query Match 15.2%; Score 34; DB 3; Length 42;  
Best Local Similarity 40.0%; Pred. No. 1.9e+03;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Search completed: February 21, 2002, 16:41:29  
Job time: 184 sec

QY 23 SNPTDDVSSGSSRSSTS 42  
| : | | | | :  
Db 1 SYSSSSASSSSSSSHSSHA 20

## RESULT 14

Q9UH60  
ID Q9UH60 PRELIMINARY; PRT; 42 AA.  
AC Q9UH60;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE CYCLIN-DEPENDENT KINASE INHIBITOR P27 (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Waltregny D., Loda M.;  
RT "Homo sapiens cyclin-dependent kinase inhibitor p27 intron."  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF213700; AAF21058.1;  
KW Kinase; Cyclin.  
FT NON\_TER 1 1  
FT 42 42  
SQ SEQUENCE 42 AA; 4461 MW; 92B374982A590CD9 CRC64;

Query Match 15.2%; Score 34; DB 4; Length 42;  
Best Local Similarity 41.2%; Pred. No. 1.9e+03;  
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 23 SNPTDDVSSGSSRS 39  
: | : : : | | : |  
Db 24 ANRTEENVSDGSPNAGS 40

## RESULT 15

Q9R508  
ID Q9R508 PRELIMINARY; PRT; 38 AA.  
AC Q9R508;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE ROLA PROTEIN (FRAGMENT).  
OS Agrobacterium rhizogenes.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=359;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96082572; PubMed=7581667;  
RA Michael T., Spena A.;  
RT "The plant oncogenes rola, B, and C from Agrobacterium rhizogenes.  
Effects on morphology, development, and hormone metabolism.";  
RL Methods Mol. Biol. 44:207-222(1995).  
SQ SEQUENCE 38 AA; 3927 MW; F90FA27E17C90AB7 CRC64;

Query Match 15.0%; Score 33.5; DB 2; Length 38;  
Best Local Similarity 36.8%; Pred. No. 2e+03;  
Matches 7; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 14 RINPDIYPSNPTD-DDVS 31  
| : | | : | : | :  
Db 8 RVSPGDVPPDQVAELDDL 26



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2002, 16:35:10 ; Search time 23 6 Seconds  
(without alignments)  
131.825 Million cell updates/sec

Title: US-08-753-851-4  
Perfect score: 223  
Sequence: 1 CRDTRYVQGEYRINPEDI.....SNPTDDVSSGSSRSSTS 42

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 267838

Minimum DB seq length: 0  
Maximum DB seq length: 42

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : A\_Geneseq\_l101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	223	100.0	42	15 AAR53476	CD44 peptide CD44-
2	173	77.6	32	15 AAR53475	CD44 peptide CD44-
3	126	56.5	28	AA15201	Surrounding sequen
4	105	47.1	21	15 AAR53477	CD44 peptide CD44-
5	96	43.0	30	15 AAR53481	CD44 peptide CD44-
6	47	21.1	38	20 AAY42816	Erythropoietin rec
7	46.5	20.9	37	21 AAB21600	Cone snail alpha-c
8	44	19.7	42	22 AAM34861	Peptide #8898 enco
9	43.5	19.5	39	19 AAM42089	BS106 sythetic po
10	43.5	19.5	39	20 AAY13467	BS106 polypeptide
11	43	19.3	20	20 AAY01162	Polypeptide fragme

12	41	18.4	38	21	AAY86368	Human gene 11-enco
13	40.5	18.2	35	22	AAB48726	Mouse liver growth
14	40	17.9	25	8	AAP71323	Phosphopeptide 4.
15	40	17.9	25	12	AAR14450	Phosphopeptide #4
16	40	17.9	25	13	AAR28428	Anticariogenic pho
17	40	17.9	25	14	AAR32930	Casein phosphopept
18	40	17.9	25	14	AAR31240	Phosphopeptide 4.
19	40	17.9	25	14	AAR32311	Phosphopeptide der
20	40	17.9	25	15	AAR47817	Sequence of casein
21	40	17.9	25	16	AAR68939	Sodium caseinate t
22	40	17.9	25	19	AAM66601	Bos alpha-s2-casei
23	40	17.9	25	21	AAB12803	Phosphopeptide T4.
24	40	17.9	35	20	AAM88473	Human secretory pe
25	40	17.9	35	21	AAY15133	Human zsig9 secret
26	40	17.9	36	21	AAB38408	Fragment of human
27	40	17.9	38	16	AAR62966	zinc finger 2 sequ
28	39.5	17.7	40	19	AAM43097	Rabbit polymeric i
29	39	17.5	24	20	AAY08793	Yeast Nup2 protein
30	39	17.5	24	20	AAY08821	Yeast YNup2 protei
31	39	17.5	41	20	AAY60433	Human normal bladd
32	37.5	16.8	31	22	AAM06219	Peptide #4901 enco
33	37	16.6	15	14	AAR39572	Sequence of a new
34	37	16.6	15	21	AAY90828	Linker amino acid
35	37	16.6	15	22	AU04949	Humanised anti-pl8
36	37	16.6	17	21	AAY67939	EPOA-hSA fusion pr
37	37	16.6	28	19	AAM66707	Peptide derived fr
38	37	16.6	31	21	AAB15670	Alpha V beta 3 rec
39	37	16.6	37	18	AAM26641	H. insolens family
40	37	16.6	38	19	AAM45595	Peptide recognised
41	37	16.6	41	20	AAY26963	Human TrkA protein
42	36.5	16.4	21	19	AAM42092	BS106 sythetic po
43	36.5	16.4	21	20	AAY13470	BS106 polypeptide
44	36.5	16.4	36	22	AAM16778	Peptide #3212 enco
45	36.5	16.4	36	22	AAM29263	Peptide #3300 enco

## ALIGNMENTS

RESULT 1  
AAR53476  
ID AAR53476 standard; peptide; 42 AA.  
XX  
AC AAR53476;  
XX  
DT 01-DEC-1994 (first entry)  
XX  
DE CD44 peptide CD44-4.  
XX  
KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autoimmune disease;  
KW immunosuppressive.  
XX  
OS Homo sapiens.  
XX  
PN W09409811-A.  
XX  
PD 11-MAY-1994.  
XX  
PF 29-OCT-1993; 93WO-0510412.  
XX  
PR 30-OCT-1992; 92US-0973339.  
XX  
(UYDU-) UNIV DUKE.  
XX  
PA Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
XX  
PI WPI; 1994-167121/20.  
XX  
PT Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis  
XX

QY 1 CRDGRYVQKGEYRINPEDIYPSNPTDDDVSS 32

```

XX DE CD44 peptide CD44-5.
XX KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;
XX KW arthritis; tumor cell metastasis; autoimmune disease;
XX KW immunosuppressive.
XX OS Homo sapiens.
XX PN WO9409811-A.
XX PD 11-MAY-1994.
XX PF 29-OCT-1993; 93WO-US10412.
XX PR 30-OCT-1992; 92US-0973339.
XX PA (UYDU-) UNIV DUKE.
XX PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;
XX DR WPI; 1994-167121/20.
XX PT Use of CD44 protein and new peptide derivs - for developing prods
XX PT for inflammation, immune-mediated tissue damage and tumour cell
XX PS metastasis
XX PS Claim 4; Page 14; 83pp; English.
XX CC The peptide can be used for treating inflammation and immune-
XX CC mediated tissue damage such as occurs in the course of autoimmune
XX CC diseases, e.g. rheumatoid arthritis. It can also be used for
XX CC determining metastatic potential or to treat or prevent tumor cell
XX CC metastasis. This peptide corresponds to AA 150-161, 170-177 and
XX CC 189-198 of the CD44
XX CC protein sequence.
XX SQ Sequence 30 AA;

Query Match 43.0%; Score 96; DB 15; Length 30;
Best Local Similarity 67.9%; Pred. No. 9.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 CRDTRYVQKGEYRINPEDIYPSNPTDD 28
DB ||||| ||||| ||||| |||||
1 crdgyrvqkge-----ypsnptdd 20

RESULT 6
AAV42816
ID AAY42816 standard; peptide; 38 AA.
XX AC AAY42816;
XX DT 19-JAN-2000 (first entry)
XX DE Erythropoietin receptor (EPOR)-binding peptide B6.
XX KW EPO; receptor; erythropoietin receptor; EPOR; immunoglobulin G; IgG1;
XX KW fusion protein; vector; cloning; synthetic peptide; biopanning;
XX KW evolved library; mimetic; proliferation; biological activity;
XX KW chronic renal failure; malignancy; cancer; chemotherapy; HIV;
XX KW rheumatoid arthritis; treatment; anaemia.
XX OS Synthetic.
XX PN WO9947151-A1.
XX PD 23-SEP-1999.
XX PF 17-MAR-1999; 99WO-US05842.
XX PR 20-MAR-1998; 98US-0078784.
XX PA (CHUS ) CHUGAI PHARM CO LTD.
XX PI McConnell SJ, Spinella DG;
XX WPI; 1999-601069/51.
XX DR N-PSDB; AA230863.
XX PT New ligands for human erythropoietin receptor, useful for treatment of
XX PT anemia
XX PS Example 7; Fig 4; 84pp; English.
XX CC This sequence represents the erythropoietin receptor (EPOR)-binding
XX CC peptide B6. This peptide was identified as being able to bind a
XX CC recombinant erythropoietin receptor via biopanning methods. The
XX CC peptide is a product of an ERB1 (erythropoietin receptor (EPOR)

```



QY 3 DGTRYVQGEYRINPEDIYPSNP---TDDDVSSG 33

RESULT 8  
AAM34861  
ID AAM34861 standard; Protein: 42 AA.

AA  
AC  
XX  
DT

AAM34861;  
17-OCT-2001 (first entry)

DE XX  
 DE XX  
 KW KW Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX XX  
 OS Homo sapiens.  
 XX XX  
 PN WO200157272-A2.  
 XX XX  
 PD 09-AUG-2001.

XX	30-JAN-2001; 2001WO-US00663,
PF	
XX	
XX	04-FEB-2000; 2000US-0180312,
PR	
PR	26-MAY-2000; 2000US-0207456

PR 30-JUN-2000; 2000US-0608408  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.  
XX

PA (MOLE-%) MOLECULAR DYNAMICS INC.  
XX  
XX

PI PENN SG, HANZEL DK, CHEN W, KAIK DK, XX

DR WPI; 2001-488897/53.  
XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -

XX  
PS  
Claim 27. SEQ ID NO 35130. 654pp. English

PS CLAIM 27; SEQ ID NO 35130; 634pp; ENGLISH: XX

The present invention relates to single exon nuclear acid probes (*SENPs*). The present invention relates to AAI131315-AAI157546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.

CC predicting, measuring and displaying gene expression in samples  
CC such probe. The probes are useful for producing a microarray in  
CC from human placenta. The probes are useful for antenatal diagnosis  
CC human genetic disorders.

XX  
CC human genetic dis

```

SQ Sequence 42 AA:
      Query Match      19.7%; Score 44; DB 22; Length 42;
      Best Local Similarity 40.7%; Pred. No. 35;
      Matches 11; Conservative 5; Mismatches 5; Indels
OY 22 PSNPPTDDVVSGSS-----SERSSTS 42

```

QY 22 PSNPTDDDVSSGSS-----SERSSTS 42 .







```

FH Key Location/Qualifiers
FT Modified-site 11 /label= OTHER
FT /note= "phosphoserine (Pse)"
FT Modified-site 12 /label= OTHER
FT /note= "Pse"
FT Modified-site 13 /label= OTHER
FT /note= "Pse"
FT Modified-site 16 /label= OTHER
FT /note= "Pse"
XX US5015628-A.
PN
XX
XX 14-MAY-1991.
PD
XX
XX 03-AUG-1990; 90US-0563798.
XX
XX 03-AUG-1990; 90US-0563798.
XX
XX (UYME-) UNIV OF MELBOURNE.
XX (VICT-) VICTORIAN DAIRY INDUSTRY.
XX
XX Reynolds EC;
XX
XX WPI; 1991-316875/43.
DR
XX
XX Novel phosphopeptides - useful for treating dental diseases,
XX rarefying diseases or diseases relating to malabsorption of minerals
XX
XX Claim 1; Column 11; 8pp; English.
XX
XX This is an example of a highly generic formula for a phosphopeptide
XX of length 5-13 amino acids. The peptides can be made synthetically
XX (e.g. chemical synthesis or genetic engineering) or they can be
XX extracted from cereals, nuts or vegetables or by fractionating a
XX digest of casein, alpha-s-casein, beta-casein or a salt of it.
XX Compositions comprising the peptide may take the form of foodstuff
XX or confectionery, dentifrices; mouthwashes and preparations for
XX topical application to teeth or gingival tissue. The peptides
XX significantly increase absorption of calcium, phosphate and iron in
XX the gut. See AAR14447-R14451.
XX
SQ Sequence 25 AA;

Query Match 17.9%; Score 40; DB 12; Length 25;
Best Local Similarity 47.1%; Pred. No. 70;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 24 NPTDDVSSGSSSSRS 40
Db 1 naneeysigsseesa 17

```

Search completed: February 21, 2002, 16:38:21  
Job time: 191 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:36:50 ; Search time 12.47 seconds  
(without alignments)  
75.793 Million cell updates/sec

Title: US-08-753-851-4

Perfect score: 223

Sequence: 1 CRDGRYVQGEYRINPEDI.....SNPTDDVSGSSRSSTS 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 145938

Minimum DB seq length: 0

Maximum DB seq length: 42

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	223	100.0	42	2	US-08-143-311B-4
2	173	77.6	32	2	US-08-143-311B-3
3	105	47.1	21	2	US-08-143-311B-5
4	96	43.0	30	2	US-08-143-311B-9
5	41	18.4	42	1	US-08-078-683A-12
6	40	17.9	25	1	US-07-748-344B-4
7	40	17.9	25	2	US-08-954-985A-4
8	39.5	17.7	40	3	US-08-856-383-6
9	39	17.5	24	4	US-09-215-221-11
10	37	16.6	15	1	US-08-133-804-8
11	37	16.6	15	1	US-08-461-386-8
12	37	16.6	15	2	US-08-461-386-8
13	37	16.6	15	2	US-08-356-786-12
14	37	16.6	15	2	US-08-665-202-3
15	37	16.6	15	3	US-08-483-749A-31
16	37	16.6	28	2	US-09-017-205-85
17	37	16.6	37	3	US-08-814-052-37
18	37	16.6	37	3	US-08-812-829-29
19	36	16.1	35	1	US-08-159-340A-42
20	36	16.1	36	2	US-08-437-607A-48
21	36	16.1	37	4	US-08-931-858E-175
22	36	16.1	38	2	US-08-726-306A-88
23	35.5	15.9	31	1	US-08-053-131-83
24	35.5	15.9	31	1	US-08-645-641-83
25	35.5	15.9	31	1	US-07-853-408B-83
26	35.5	15.9	31	2	US-08-096-762-83
27	35.5	15.9	31	2	US-08-308-865-83

28 35.5 15.9 31 4 US-09-042-353-280 Sequence 280, App  
29 35.5 15.9 31 4 US-08-758-417A-128 Sequence 128, App  
30 35.5 15.9 31 5 PCT-US92-10983-83 Sequence 83, Appl  
31 35.5 15.9 36 4 US-09-082-279B-828 Sequence 828, App  
32 35 15.7 12 4 US-08-938-595-5 Sequence 5, Appl  
33 35 15.7 12 4 US-08-727-153-5 Sequence 5, Appl  
34 35 15.7 14 2 US-08-373-190-3 Sequence 3, Appl  
35 35 15.7 14 2 US-08-438-190A-3 Sequence 3, Appl  
36 35 15.7 14 3 US-08-350-215-3 Sequence 3, Appl  
37 35 15.7 14 3 US-09-287-145A-3 Sequence 3, Appl  
38 35 15.7 15 2 US-08-373-190-4 Sequence 4, Appl  
39 35 15.7 15 2 US-08-438-190A-4 Sequence 4, Appl  
40 35 15.7 15 3 US-08-350-215-4 Sequence 4, Appl  
41 35 15.7 15 3 US-09-287-145A-4 Sequence 4, Appl  
42 35 15.7 24 1 US-08-446-692-8 Sequence 8, Appl  
43 35 15.7 24 2 US-08-488-351A-8 Sequence 8, Appl  
44 35 15.7 24 3 US-09-100-409A-44 Sequence 44, Appl  
45 35 15.7 24 5 PCT-US95-13841-11 Sequence 11, Appl

#### ALIGNMENTS

RESULT 1  
US-08-143-311B-4  
; Sequence 4, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-4

Query Match 100.0%; Score 223; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 4.2e-24;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRDGRYVKGGRINPEDIYPSNPTDDVSSGSSRSSTS 42  
|||||  
Db 1 CRDGRYVKGGRINPEDIYPSNPTDDVSSGSSRSSTS 42

RESULT 2  
US-08-143-311B-3  
; Sequence 3, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-3

Query Match 77.6%; Score 173; DB 2; Length 32;  
Best Local Similarity 96.9%; Pred. No. 2.2e-17;  
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRDGRYVKGGRINPEDIYPSNPTDDVSS 32  
|||||  
Db 1 CRDGRYVKGGRINPEDIYPSNPTDDVSS 32

RESULT 3

US-08-143-311B-5  
; Sequence 5, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-5

Query Match 47.1%; Score 105; DB 2; Length 21;  
Best Local Similarity 95.0%; Pred. No. 2.9e-08;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 YRNPEDIYPSNPTDDVSS 32  
|||  
Db 2 YRNPEDIYPSNPTDDVSS 21

RESULT 4  
US-08-143-311B-9  
; Sequence 9, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/143.311B  
FILING DATE: 29-OCT-1993  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973.339  
FILING DATE: 30-OCT-1992  
CLASSIFICATION: 436  
APPLICATION NUMBER: 07/669.730  
FILING DATE: 15-MAR-1991  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-143-311B-9

Query Match 43.0%; Score 96; DB 2; Length 30;  
Best Local Similarity 67.9%; Pred. No. 7.8e-07;  
Matches 19; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 CRDGRYVQKGEYRNPEDYPSNPTDD 28  
|||||  
DB 1 CRDGRYVQKGE-----YPSNPTDD 20  
|||||

RESULT 5  
US-08-078-683A-12  
Sequence 12, Application US/08078683A  
Patent No. 5486599  
GENERAL INFORMATION:  
APPLICANT: Saunders, Scott  
APPLICANT: Bernfield, Merton  
APPLICANT: Kato, Masato  
TITLE OF INVENTION: Construction and Use of Synthetic  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII (text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/078.683A

FILING DATE: 17-JUN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CME-062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: Internal  
US-08-078-683A-12

Query Match 18.4%; Score 41; DB 1; Length 42;  
Best Local Similarity 62.5%; Pred. No. 43;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 21 YPSNPTDDVSSGSS 36  
||| : ||| |||||  
DB 11 YPVD--DDYSSGSGS 24

RESULT 6  
US-07-748-344B-4  
Sequence 4, Application US/07748344B  
Patent No. 5227154  
GENERAL INFORMATION:  
APPLICANT: REYNOLDS, ERIC CHARLES  
TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE  
TREATMENT OF DENTAL CALCULUS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDRUS, SCALES, STARKE & SAWALL  
STREET: 100 EAST WISCONSIN AVE., SUITE 1100  
CITY: MILWAUKEE  
STATE: WISCONSIN  
COUNTRY: USA  
ZIP: 53202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/748,344B  
FILING DATE: 19910822  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: SARA, CHARLES S  
REGISTRATION NUMBER: 30,492  
REFERENCE/DOCKET NUMBER: C.8493-87  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 255-2022  
TELEFAX: (608) 255-2182  
TELEX: 26832 ANDSTARK  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
FEATURE:  
NAME/KEY: Phosphoserine



OTHER INFORMATION: Post-translational

Matches 12; Conservative 6; Mismatches

Query Match 17.7%; Score 39.5; DB 3; Length 40;  
Best Local Similarity 34.3%; Pred. No. 65;  
Matches 12: Conservative 6; Mismatches 14: Indels 3; Gaps 1;

```

QY      8 VQGEYRINPDDYPNPTDDDVSSGSSSRSTS 42
    II I :I   III : I |III| : I
Db      6 VQSAE---DPASGDPAAGSRSRVDSGGSEEGGSS 37

RESULT          9
US-09-215-221-11
; Sequence 11, Application US/09215221
; Patent No. 6265562
; GENERAL INFORMATION:
; APPLICANT: EILERS, MARTIN
; APPLICANT: BUERGIN, ANDREA
; TITLE OF INVENTION: NUCLEIC ACID CONSTRUCTS WHOSE ACTIVITY IS AFFECTED BY
; TITLE OF INVENTION: INHIBITORS OF CYCLIN-DEPENDANT KINASES AND USES THEREOF
; FILE REFERENCE: 026083/0192
; CURRENT APPLICATION NUMBER: US/09/215,221
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 197 56 975.7
; PRIOR FILING DATE: 1997-12-20
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-215-221-11

Query Match           17.5%; Score 39; DB 4; Length 24;
Best Local Similarity 58.3%; Pred. No. 40;
Matches              7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      19 DIYPNPTDDDV 30
    : I | I : III
Db      13 ETYDINESDDV 24

RESULT         10
US-08-133-804-8
; Sequence 8, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
```

Query Match 16.6%; Score 37; DB 1; Length 15;  
Best Local Similarity 75.0%; Pred. No. 42;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 31 SSGSSSSRSSTS 42  
Db 3 SSGSSSSGSSSS 14  
||||| ||:|

RESULT 12  
US-08-461-386-8  
; Sequence 8, Application US/08461386  
; Patent No. 5837846  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.  
; APPLICANT: Oppermann, Hermann  
; APPLICANT: Ring, David B.  
; APPLICANT: Houston, L. L.  
; TITLE OF INVENTION: Biosynthetic Binding Proteins For  
; TITLE OF INVENTION: Imaging  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,386  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kelley, Robin D.  
; REGISTRATION NUMBER: 34,637  
; REFERENCE/DOCKET NUMBER: 2054/22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-248-7477  
; TELEFAX: 617-248-7100  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..15  
; OTHER INFORMATION: /note= "LINKER 2"  
US-08-461-386-8

Query Match 16.6%; Score 37; DB 2; Length 15;  
Best Local Similarity 75.0%; Pred. No. 42;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 31 SSGSSSSRSSTS 42  
Db 3 SSGSSSSGSSSS 14  
||||| ||:|

RESULT 13  
US-08-356-786-12  
; Sequence 12, Application US/08356786  
; Patent No. 5877305  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.

; APPLICANT: Oppermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
; TITLE OF INVENTION: Marker  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-356-786-12

Query Match 16.6%; Score 37; DB 2; Length 15;  
Best Local Similarity 75.0%; Pred. No. 42;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 31 SSGSSSSRSSTS 42  
Db 3 SSGSSSSGSSSS 14  
||||| ||:|

RESULT 14  
US-08-665-202-3  
; Sequence 3, Application US/08665202  
; Patent No. 5977322  
; GENERAL INFORMATION:  
; APPLICANT: Marks, James D.  
; APPLICANT: Schier, Robert  
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to  
; TITLE OF INVENTION: Tumor Antigens  
; NUMBER OF SEQUENCES: 141  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

\* APPLICATION NUMBER: US/08/665,202  
FILING DATE: 13-JUN-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000,238  
FILING DATE: 14-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000,250  
FILING DATE: 15-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 02307E-061410  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-665-202-3

Query Match 16.6%; Score 37; DB 2; Length 15;  
Best Local Similarity 75.0%; Pred. No. 42;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 31 SSGSSSSRSSTS 42  
Db 3 SSGSSSSGSSSS 14

## RESULT 15

US-08-483-749A-31  
Sequence 31, Application US/08483749A  
Patent No. 6054561  
GENERAL INFORMATION:  
APPLICANT: RING, DAVID B.  
TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY  
TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHIRON CORPORATION  
STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097  
CITY: EMERYVILLE  
STATE: CA  
COUNTRY: USA  
ZIP: 94662-8097  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,749A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: SAVERIDE, PAUL B.  
REGISTRATION NUMBER: 36,914  
REFERENCE/DOCKET NUMBER: 0508.008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2585  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-483-749A-31

Query Match 16.6%; Score 37; DB 3; Length 15;  
Best Local Similarity 75.0%; Pred. No. 42;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 31 SSGSSSSRSSTS 42  
Db 3 SSGSSSSGSSSS 14

Search completed: February 21, 2002, 16:38:40  
Job time: 110 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:41:06 ; Search time 33.6 Seconds  
(without alignments)  
47.609 Million cell updates/sec

Title: US-08-753-851-5  
Perfect score: 120  
Sequence: 1 CYRTNPEDIYPSNPTDDVSS 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4046

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	33	27.5	19	2 A41299	T-cell receptor al
2	29.5	24.6	19	2 A61377	endometrial secret
3	28	23.3	18	2 S24780	protein-tyrosine k
4	27	22.5	15	2 S27248	pseudogerm - whe
5	26	21.7	18	2 A39040	calsequestrin, car
6	26	21.7	18	2 A32917	protein phosphatas
7	26	21.7	20	2 S77983	cytochrome-c oxida
8	25.5	21.2	16	2 A23992	melatin-omochrome
9	25	20.8	15	2 G49732	NADH dehydrogenase
10	25	20.8	18	2 A54651	insulin-like growt
11	25	20.8	20	2 S72501	protein kinase C i
12	25	20.8	21	2 S78416	ribosomal protein
13	24.5	20.4	20	2 S00492	hemocyanin chain I
14	24	20.0	8	2 A61328	trypsin (EC 3.4.21
15	24	20.0	11	2 D61033	ranatachykinin D -
16	24	20.0	13	2 G37266	Ig heavy chain C r
17	24	20.0	15	2 PQ0545	capsid protein VPI
18	24	20.0	15	2 PA0088	protein QF200051 -
19	24	20.0	16	2 D58501	26K kidney and gal
20	24	20.0	19	2 PH1330	Ig heavy chain DJ
21	24	20.0	20	2 S09022	carboxylesterase (
22	24	20.0	20	2 S50743	proteinase inhibit
23	23	19.2	12	2 E44787	callimRFamide 1 -
24	23	19.2	13	2 S21152	tryptophyllin-rela
25	23	18.2	13	2 A05174	tryptophyllin-13 -
26	23	19.2	15	2 B59045	alpha-conotoxin Au
27	23	19.2	15	2 A48372	benzoyl-CoA ligase
28	23	19.2	16	2 A59045	alpha-conotoxin Au
29	23	19.2	16	2 A49226	major outer membra

30 23 19.2 19 2 S65434 manganese peroxida  
31 23 19.2 20 2 S50022 trypsin-like prote  
32 23 19.2 20 2 T44453 acetyl-CoA synthet  
33 23 19.2 20 2 S44465 pyrrhocoricin - Py  
34 23 19.2 21 2 S05031 cytochrome b559 co  
35 22.5 18.8 16 2 PS0210 28K protein 4209 -  
36 22 18.3 11 2 PC2330 cycloinnuloooligosac  
37 22 18.3 11 2 S43626 cytochrome-c oxida  
38 22 18.3 14 2 PL0152 metal-binding prot  
39 22 18.3 15 2 A56049 urinary tract ston  
40 22 18.3 16 2 C59045 alpha-conotoxin Au  
41 22 18.3 17 2 C37396 pollen allergen Fe  
42 22 18.3 18 1 A58589 alpha-conotoxin EI  
43 22 18.3 18 2 PN0149 beta-Gliadine 13 -  
44 22 18.3 18 2 S23971 alpha-macroglobuli  
45 22 18.3 18 2 A56798 dermatan sulfate p

ALIGNMENTS

RESULT 1

A41299  
T-cell receptor alpha chain precursor V region (17.2) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 23-Jul-1999  
C:Accession: A41299  
R:Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayl, G  
Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991  
A:Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheum  
A:Reference number: A41299; MUID:92020887  
A:Accession: A41299  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-19 <UEM>  
A:Cross-references: GB:S57433; NID:g236318; PIDN:AAB19556.1; PID:g236319  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor

Query Match 27.5%; Score 33; DB 2; Length 19;  
Best Local Similarity 58.3%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 8 DIYPSNPTDDDV 19  
||| | | | |  
Db 4 DIVPSQPCDSAV 15

RESULT 2

A61377  
endometrial secretory protein - sheep (fragment)  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 17-Mar-1999  
C:Accession: A61377  
R:Vallet, J.L.; Barker, P.J.; Lammung, G.E.; Skinner, N.; Huskisson, N.S.  
J. Endocrinol. 130, R1-R4, 1991  
A:Title: A low molecular weight endometrial secretory protein which is increased by o  
A:Reference number: A61377; MUID:92013712  
A:Accession: A61377  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-19 <VAL>

Query Match 24.6%; Score 29.5; DB 2; Length 19;  
Best Local Similarity 46.2%; Pred. No. 3.1e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 6 PE-DIYPSNPTDD 17  
|| | | | | |  
Db 5 PEVQVSSDDPPED 17

## RESULT 3

S24780  
 protein-tyrosine kinase (EC 2.7.1.112) lck - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 19-Feb-1994 #sequence\_revision 26-Jul-1996 #text\_change 04-Mar-2000  
 C:Accession: I58370; S24780  
 R:Shin, S.; Steffen, D.L.  
 Oncogene 8, 141-149, 1993  
 A:Title: Frequent activation of the lck gene by promoter insertion and aberrant splicing  
 A:Reference number: I58370; MUID:93141260  
 A:Accession: I58370  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-18 <RES>  
 A:Cross-references: EMBL:215029; NID:g56567; PIDN:CAA78748.1; PID:g56568  
 C:Genetics:  
 A:Gene: lck  
 C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology  
 C:Keywords: ATP; blocked amino end; lipoprotein; myristylation; phosphorylation; phosphotransferase; thio  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:3,5/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 23.3%; Score 28; DB 2; Length 18;  
 Best Local Similarity 83.3%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 SNTDD 17  
 IIII  
 Db 7 SNPEDD 12

## RESULT 4

S27248  
 pseudogerm - wheat  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Aug-1997  
 C:Accession: S27248  
 R:Lane, B.G.; Cumming, A.C.; Fregeau, J.; Carpita, N.C.; Harkman, W.J.; Bernier, F.; Drat  
 Eur. J. Biochem. 209, 961-969, 1992  
 A:Title: Germ in isoforms are discrete temporal markers of wheat development. Pseudogerm  
 ated embryos, it is incorporated into cell walls.  
 A:Reference number: S27247; MUID:93049354  
 A:Accession: S27248  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <LAN>

Query Match 22.5%; Score 27; DB 2; Length 15;  
 Best Local Similarity 80.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 YPSNP 14  
 IIII  
 Db 4 YPXNP 8

## RESULT 5

A39040  
 calsequestrin, cardiac muscle - dog (fragments)  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 23-Feb-1997  
 C:Accession: A39040  
 R:Calal, S.E.; Jones, L.R.  
 J. Biol. Chem. 266, 391-398, 1991  
 A:Title: Phosphorylation of cardiac and skeletal muscle calsequestrin isoforms by casein  
 A:Reference number: A39040; MUID:91093153  
 A:Accession: A39040  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-18 <CAL>

C:Keywords: cardiac muscle; heart; phosphoprotein

Query Match 21.7%; Score 26; DB 2; Length 18;  
 Best Local Similarity 41.7%; Pred. No. 9.8e+02;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 EDIYPSNPTDDD 18  
 : I I : I I  
 Db 5 DDDLPTINTEDD 16

## RESULT 6

A32917  
 protein phosphatase - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 08-Dec-1989 #sequence\_revision 08-Dec-1989 #text\_change 30-Sep-1993  
 C:Accession: A32917  
 R:Tamura, S.; Yasui, A.; Tsulki, S.  
 Biochem. Biophys. Res. Commun. 163, 131-136, 1989  
 A:Title: Expression of rat protein phosphatase 2C (1A) in Escherichia coli.  
 A:Reference number: A32917; MUID:89374212  
 A:Accession: A32917  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-18 <TAM>

Query Match 21.7%; Score 26; DB 2; Length 18;  
 Best Local Similarity 44.4%; Pred. No. 9.8e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 11 PSNPTDDDV 19  
 I : I : I  
 Db 7 PGDPLEDDV 15

## RESULT 7

S77983  
 cytochrome-c oxidase (EC 1.9.3.1) chain Vb - bigeye tuna (fragment)  
 C:Species: Thunnus obesus (bigeye tuna)  
 C:Date: 17-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 30-Jan-1998  
 C:Accession: S77983  
 R:Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.  
 submitted to the Protein Sequence Database, June 1997  
 A:Reference number: S77980  
 A:Accession: S77983  
 A:Molecule type: protein  
 A:Residues: 1-20 <ARN>  
 A:Experimental source: heart; liver  
 C:Genetics:  
 A:Genome: nuclear  
 C:Function:  
 A:Pathway: oxidative phosphorylation; respiratory chain  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 21.7%; Score 26; DB 2; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 PTDDD 18  
 IIII  
 Db 7 PTDD 11

## RESULT 8

A23992  
 melatin-omochrome-stimulating hormone III - silkworm (fragment)  
 N:Alternate names: melanization and reddish coloration hormone III; MRCH III  
 C:Species: Bombyx mori (silkworm)  
 C:Date: 30-Jan-1988 #sequence\_revision 30-Jan-1988 #text\_change 18-Jun-1993  
 C:Accession: A23992

R:Matsumoto, S.; Isogai, A.; Suzuki, A.  
 FEBS Lett. 189, 115-118, 1985  
 A:Title: N-terminal amino acid sequence of an insect neurohormone, melanization and redd  
 A:Reference number: A23992; MUID:85285612  
 A:Accession: A23992  
 A:Molecule type: protein  
 A:Residues: 1-16 <MAT>  
 C:Keywords: hormone

Query Match 21.2%; Score 25.5; DB 2; Length 16;  
 Best Local Similarity 38.5%; Pred. No. 1e+03;  
 Matches 5; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 7 EDIYPSNPTDDV 19  
 ||: | : | : | :  
 DB 4 EDM-PATPADQEM 15

## RESULT 9

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) 20K chain - potato mitochondrion (fragment)  
 N:Alternate names: complex I hydrogenase 20K chain; NADH-ubiquinone oxidoreductase 20K c  
 C:Species: mitochondrion Solanum tuberosum (potato)  
 C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 07-May-1999  
 C:Accession: G49732  
 R:Herz, U.; Schroeder, W.; Liddell, A.; Leaver, C.J.; Brennicke, A.; Grohmann, L.  
 J. Biol. Chem. 269, 2263-2269, 1994  
 A:Title: Purification of the NADH:ubiquinone oxidoreductase (complex I) of the respirato  
 A:Reference number: A49732; MUID:94124587  
 A:Accession: G49732  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <HER>  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Keywords: electron transfer; hydrogen ion transport; mitochondrial inner membrane; mit

Query Match 20.8%; Score 25; DB 2; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 12 SNPTDD---DVSS 21  
 || ||: |||:  
 DB 2 SNATDETMLKDVSA 15

## RESULT 10

insulin-like growth factor-binding protein 22K form - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 13-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 11-Apr-1995  
 C:Accession: A54651  
 R:Rognani, M.; Segovia, B.; Whitechurch, O.; Binoux, M.  
 Growth Regul. 1, 125-130, 1991  
 A:Title: Purification from human cerebrospinal fluid of insulin-like growth factor bindi  
 A:Reference number: A54651; MUID:93091816  
 A:Accession: A54651  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-18 <ROG>  
 A:Experimental source: cerebrospinal fluid  
 A:Note: sequence extracted from NCHI backbone (NCBIP:121074)

Query Match 20.8%; Score 25; DB 2; Length 18;  
 Best Local Similarity 26.7%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 EDIYPSNPTDDVSS 21  
 : | | : |  
 DB 1 DSFVPEPSPDEKALS 15

## RESULT 11

protein kinase C inhibitor - human (fragment)  
 N:Alternate names: histidine triad nucleotide-binding protein  
 C:Species: Homo sapiens (man)  
 C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
 C:Accession: S72501; S62623  
 R:Maines, M.D.; Trakshel, G.M.  
 Arch. Biochem. Biophys. 300, 320-326, 1993  
 A:Title: Purification and characterization of human biliverdin reductase.  
 A:Reference number: S29736; MUID:93143333  
 A:Accession: S72501  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAI>  
 A:Note: this protein was identified as biliverdin reductase; the identification is qu  
 R:Maines, M.D.; Polevoda, B.V.; Huang, T.J.; McCoubrey Jr., W.K.  
 Eur. J. Biochem. 235, 372-381, 1996

A:Title: Human biliverdin IX-alpha reductase is a zinc-metalloprotein. Characterizati  
 A:Reference number: S62622; MUID:96202961  
 A:Accession: S62623  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAW>  
 C:Superfamily: protein kinase C inhibitor; histidine triad homology  
 C:Keywords: homodimer; protein kinase inhibitor; zinc

Query Match 20.8%; Score 25; DB 2; Length 20;  
 Best Local Similarity 62.5%; Pred. No. 1.5e+03;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 DIYPSNPT 15  
 || | ||  
 DB 6 DISPOAPT 13

## RESULT 12

ribosomal protein RL28, mitochondrial [validated] - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jul-2000  
 C:Accession: S78416  
 R:Goldschmidt-Reisin, S.; Graack, H.R.  
 submitted to the Protein Sequence Database, February 1998  
 A:Reference number: S78411  
 A:Accession: S78416  
 A:Molecule type: protein  
 A:Residues: 1-21 <GOL>  
 A:Note: the protein is designated as mitochondrial ribosomal protein L28  
 C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 20.8%; Score 25; DB 2; Length 21;  
 Best Local Similarity 23.1%; Pred. No. 1.6e+03;  
 Matches 3; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 6 PEDIYPSNPTDDD 18  
 | : : | : | :  
 DB 8 PDKVQFQRPEDHE 20

## RESULT 13

hemocyanin chain Ia - Japanese spiny lobster (fragment)  
 C:Species: Panulirus japonicus (Japanese spiny lobster)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 01-Aug-1997  
 C:Accession: S00492  
 R:Makino, N.; Kimura, S.  
 Eur. J. Biochem. 173, 423-430, 1988  
 A:Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.  
 A:Reference number: S00492; MUID:88196131  
 A:Accession: S00492



A:Molecule type: protein  
A:Residues: 1-20 <MAK>  
C:Superfamily: hemocyanin  
C:Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 20.4%; Score 24.5; DB 2; Length 20;  
Best Local Similarity 36.8%; Pred. No. 1.1e+03;  
Matches 7; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

QY 8 DIYPSNPTD-----DDV 19  
| | : | |  
Db 2 DXQPGDSTDKLLAQKQDDV 20

## RESULT 14

A61328  
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)  
C:Species: Balaenoptera acutorostrata (minke whale, lesser rorqual)  
C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 07-May-1999  
C:Accession: A61328  
R:Brictoux-Gregoire, S.; Schyns, R.; Florkin, M.; Emmens, M.; Welling, G.W.; Belintema, J.  
Biochim. Biophys. Acta 386, 244-255, 1975  
A:Title: N-terminal amino acid sequence of trypsinogen from the lesser rorqual, Balaenoptera acutorostrata

A:Reference number: A61328; MUID:75146765

A:Accession: A61328

A:Molecule type: protein

A:Residues: 1-8 <BRI>

C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen  
F:1-8/Domain: activation peptide #status experimental <APT>

Query Match 20.0%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 PTDDD 18  
| | | |  
Db 2 PIDDD 6

## RESULT 15

D61033  
ranatachykinin D - bullfrog  
C:Species: Rana catesbeiana (bullfrog)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Aug-2000  
C:Accession: D61033; JE0429  
R:Kangawa, K.; Kozawa, H.; Hino, J.; Minamino, N.; Matsuo, H.  
Regul. Pept. 42(Suppl.1), S12, 1992  
A:Title: Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and intestine  
A:Reference number: A61033  
A:Accession: D61033  
A:Molecule type: protein  
A:Residues: 1-11 <KAN>  
R:Kozawa, H.; Hino, J.; Minamino, N.; Kangawa, K.; Matsuo, H.  
Biochem. Biophys. Res. Commun. 177, 588-595, 1991  
A:Title: Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and intestine

A:Reference number: JE0426; MUID:91254337  
A:Accession: JE0429  
A:Molecule type: protein  
A:Residues: 1-11 <KOZ>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; neuropeptide  
F:11/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 20.0%; Score 24; DB 2; Length 11;  
Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 RTNPEDIY 10  
| | | | |

Db 1 KPNPERFY 8

Search completed: February 21, 2002, 16:44:15  
Job time: 189 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:43:41 ; Search time 20.15 seconds

(without alignments)  
38.211 Million cell updates/sec

Title: US-08-753-851-5

Perfect score: 120

Sequence: 1 CYRTNPEDIYPSNPTDDVSS 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1160

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	25.8	15	1 CXA2_CONAL	P56640 conus aulic
2	28	23.3	17	1 LCK_RAT	Q01621 rattus norv
3	27	22.5	19	1 HMD_METWO	P32441 methanobact
4	26	21.7	20	1 COXB_THUOB	P80374 thunnus obe
5	25	20.8	11	1 RRPL_CHAV	P13179 chandipura
6	25	20.8	15	1 NUD3_SOLTU	P80263 solanum tub
7	24	20.0	11	1 TKNA_RANRI	P29207 rana ridibu
8	24	20.0	11	1 TRND_RANCA	P22691 rana catesb
9	24	20.0	16	1 TRYP_FELCA	P81071 felis silve
10	23	19.2	10	1 UPAP2_HUMAN	P30088 homo sapien
11	23	19.2	12	1 FARI_CALVO	P41869 calliphora
12	23	19.2	13	1 TPI3_PHYRO	P04096 phyllomedus
13	23	19.2	13	1 UNO2_PINPS	P81667 pinus pinas
14	23	18.2	13	1 UPS1_OPEIN	P82036 uperoleia i
15	23	19.2	16	1 CXA1_CONAL	P56639 conus aulic
16	23	19.2	20	1 FIBB_FELCA	P14469 felis silve
17	23	19.2	20	1 PYRR_PYRAP	P37362 pyrrhocoris
18	23	19.2	21	1 PSBF_SYNVU	P12239 synecococc
19	22	18.3	16	1 CXA3_CONAL	P56641 conus aulic
20	22	18.3	18	1 A2M_OCTVU	P30800 octopus vul
21	22	18.3	18	1 CXA1_CONER	P50982 conus ermin
22	22	18.3	20	1 COXF_ONCMW	P80329 oncorhynch
23	22	18.3	21	1 CRPA_TREPH	P56738 treponema p
24	21	17.5	13	1 ACT7_SOYBN	P15987 glycine max
25	21	17.5	14	1 MAR1_ALTSP	P29399 alteromonas
26	21	17.5	14	1 RECF_SALTY	P28355 salmonella
27	21	17.5	15	1 FGF1_CANFA	P18651 canis fami
28	21	17.5	16	1 CXA1_CONEP	P56638 conus epic
29	21	17.5	16	1 CXAA_CONPE	P50984 conus penna
30	21	17.5	16	1 CXAB_CONPE	P50985 conus penna
31	21	17.5	16	1 FOR2_MYRGU	P81437 myrmecia gu
32	21	17.5	17	1 ACT6_SOYBN	P15986 glycine max
33	21	17.5	17	1 TAC1_TACGI	P23684 tachypleus

34 21 17.5 17 1 TAC3\_TACGI P18252 tachypleus  
35 21 17.5 18 1 PHL\_PERAM P82694 periplaneta  
36 21 17.5 18 1 PPM1\_LIMPO P14215 limulus pol  
37 21 17.5 19 1 PPM2\_LIMPO P14216 limulus pol  
38 21 17.5 19 1 DHAB\_COMTE P80704 comamonas t  
39 21 17.5 21 1 FIBB\_BISBO P14466 bison bonas  
40 21 17.5 21 1 FIBB\_BUBBU P14467 bubalus bub  
41 21 17.5 21 1 FIBB\_SYNCA P14481 syncerus ca  
42 20.5 17.1 16 1 MK2B\_PALPR P80410 palomona pr  
43 20 16.7 6 1 VP19\_HSVIK P23210 herpes simp  
44 20 16.7 9 1 ISOT\_CYPKA P42993 cyprinus ca  
45 20 16.7 9 1 OXYT\_RAJCL P42994 raja clavav

#### ALIGNMENTS

RESULT 1  
CXA2\_CONAL STANDARD; PRT; 15 AA.  
AC P56640;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE ALPHA-CONOTOXIN AUIB.  
OS Conus aulicus (Court cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=89437;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Venom;  
RX MEDLINE=99003392; PubMed=9786965;  
RA Luo S., Kulak J.M., Cartier G.E., Jacobsen R.B., Yoshikami D.,  
RA Olivera B.M., McIntosh J.M.;  
RT 'Alpha-conotoxin AUIB selectively blocks alpha3 beta4 nicotinic  
acetylcholine receptors and nicotine-evoked norepinephrine release.';  
RL J. Neurosci. 18:8571-8579(1998).  
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC  
ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-4 SUBUNITS.  
CC -!- MASS SPECTROMETRY: MW=1572.5; METHOD=ELECTROSPRAY.  
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
KW Venom.  
FT DISULFID 2 8  
FT DISULFID 3 15  
FT MOD\_RES 15 15  
SQ SEQUENCE 15 AA; 1578 MW; 84EFE95FDC700155 CRC64;  
AMIDATION.

Query Match 25.88; Score 31; DB 1; Length 15;  
Best Local Similarity 57.18; Pred. No. 62;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CYRTNPE 7  
I: III;  
Db 8 CFATNPD 14

RESULT 2  
LCK\_RAT STANDARD; PRT; 17 AA.  
ID LCK\_RAT  
AC Q01621;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE LCK (EC 2.7.1.112) (P56-LCK)  
(FRAGMENT).  
GN LCK.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISCHER;  
 RX MEDLINE=93141260; PubMed=8423992;  
 RA Shin S., Steffen D.L.;  
 RT "Frequent activation of the lck gene by promoter insertion and  
 aberrant splicing in murine leukemia virus-induced rat lymphomas."  
 RL Oncogene 8:141-149(1993).  
 CC -!- FUNCTION: MAY PARTICIPATE IN ANTIGEN-INDUCED T-CELL ACTIVATION.  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER  
 CC CD4 OR CD8.  
 CC -!- DISEASE: PROVIRAL INSERTION UPSTREAM OF THE LCK GENE CAUSES  
 CC OVEREXPRESSION, LEADING TO THE DEVELOPMENT OF THYMIC LYMPHOMA.  
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN. BELONGS TO THE SRC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z15029; CAA78748.1; -;  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001245; Tyr\_kin.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; PARTIAL.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; PARTIAL.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; PARTIAL.  
 DR PROTO-oncogene; Tyrosine-protein kinase; Phosphorylation; Transferase;  
 KW ATP-binding; Myristate; Palmitate; Lipoprotein.  
 FT INIT\_MET 0 0 PROBABLE.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 FT LIPID 2 2 PALMITATE (BY SIMILARITY).  
 FT LIPID 4 4 PALMITATE (BY SIMILARITY).  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1914 MW; 298848FBF0EB776 CRC64;

Query Match 23.3%; Score 28; DB 1; Length 17;  
 Best Local Similarity 83.3%; Pred. No. 2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 12 SNPTDD 17  
 III II  
 DB 6 SNPEDD 11

RESULT 3  
 HMD\_METWO  
 ID HMD\_METWO STANDARD; PRT; 19 AA.  
 AC P324A1;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE H(2)-FORMING N5,N10-METHYLENETETRAHYDROMETHANOPTERIN DEHYDROGENASE  
 DE (EC 1.5.99.11) (H2-DEPENDENT METHYLENE-H4MPT DEHYDROGENASE)  
 DE (FRAGMENT).  
 GN HMD.  
 OS Methanobacterium wolfei.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145261;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92394151; PubMed=1521540;  
 RA Zilngibl C., van Dongen W., Schwoerer B., von Buenau R.,  
 RA Richter M., Klein A., Thauer R.K.;  
 RT "H2-forming methylenetetrahydromethanopterin dehydrogenase, a novel

RT type of hydrogenase without iron-sulfur clusters in methanogenic  
 RT archaea.";  
 RL Eur. J. Biochem. 208:511-520(1992).  
 CC -!- CATALYTIC ACTIVITY: N(5),N(10)-METHYLENETETRAHYDROMETHANOPTERIN +  
 CC REDUCED COENZYME F(420) = 5-METHYL-5,6,7,8-TETRAHYDROMETHANOPTERIN  
 CC + COENZYME F(420).  
 CC -!- COFACTOR: ZINC (BY SIMILARITY).  
 CC -!- PATHWAY: INVOLVED IN METHANOGENESIS.  
 CC -!- SUBUNIT: HOMODIMER.  
 KW Oxidoreductase; Methanogenesis; Zinc.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 19 AA; 1911 MW; 0C17E9D7BF1F97C9 CRC64;

Query Match 22.5%; Score 27; DB 1; Length 19;  
 Best Local Similarity 80.0%; Pred. No. 3.2e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CYRTN 5  
 III:  
 DB 10 CYRTD 14

RESULT 4  
 COXB\_THUOB  
 ID COXB\_THUOB STANDARD; PRT; 20 AA.  
 AC P80974;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE VB (EC 1.9.3.1) (FRAGMENT).  
 OS Thunnus obesus (Bigeye tuna).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;  
 OC Scombridae; Thunnus.  
 OX NCBI\_TaxID=8241;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Heart, and Liver;  
 RX MEDLINE=97454291; PubMed=9310366;  
 RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,  
 RA Kadenbach B.;  
 RT "The subunit structure of cytochrome-c oxidase from tuna heart and  
 RT liver";  
 RL Eur. J. Biochem. 248:99-103(1997).  
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
 CC MITOCHONDRIAL ELECTRON TRANSPORT.  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.  
 DR InterPro: IPR002124; COX5B.  
 DR PROSITE: PS00848; COX5B; PARTIAL.  
 KW Oxidoreductase; Inner membrane; Mitochondrion.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2158 MW; 76F1473E1F392BD7 CRC64;

Query Match 21.7%; Score 26; DB 1; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 4.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 PTDDD 18  
 IIII:  
 DB 7 PTDD 11

RESULT 5  
 RRPL\_CHAV  
 ID RRPL\_CHAV STANDARD; PRT; 11 AA.  
 AC P13179;

DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE RNA POLYMERASE BETA SUBUNIT (EC 2.7.7.48) (LARGE STRUCTURAL PROTEIN)  
 DE (L PROTEIN) (FRAGMENT).  
 GN L.  
 OS Chandipura virus (strain 1653514).  
 OC Viruses: ssRNA negative-strand viruses; Mononegavirales;  
 OC Rhabdoviridae: Vesiculovirus.  
 OX NCBI\_TaxID=11273;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89299473; PubMed=2741347;  
 RA Masters P.S., Bhella R.S., Butcher M., Patel B., Ghosh H.P.,  
 RA Banerjee A.K.;  
 RT "Structure and expression of the glycoprotein gene of Chandipura  
 RT virus.";  
 RL Virology 171:285-290(1989).  
 CC -!- FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE  
 CC POLYMERASE. IT MAY FUNCTION IN RNA SYNTHESIS, CAPPING, AS WELL AS  
 CC METHYLATION OF CAPS, AND POLY(A) SYNTHESIS.  
 CC -!- SUBUNIT: THOUGHT TO FORM A TRANSCRIPTION COMPLEX WITH THE  
 CC NUCLEOCAPSID (N) PROTEIN.  
 CC -!- SIMILARITY: WITH THE L PROTEIN OF OTHER RHABDOVIRUSES AND  
 CC PARAMYXOVIRUSES.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J04350; AAA42917.1;  
 KW Transferase; RNA-directed RNA polymerase.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1189 MW; 0335D6E3AAB2D764 CRC64;  
 Query Match 20.8%; Score 25; DB 1; Length 11;  
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 13 NPTDD 17  
 DB 4 NPVDD 8  
 RESULT 6  
 ID NUO3\_SOLUTU STANDARD; PRT; 15 AA.  
 AC P80263;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE 20 KDA SUBUNIT (EC 1.6.5.3)  
 DE (EC 1.6.99.3) (COMPLEX I-20KD) (CI-20KD) (FRAGMENT).  
 OS Solanum tuberosum (Potato)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=CV. BINTJE; TISSUE=Tuber;  
 RX MEDLINE=94124587; PubMed=8294484;  
 RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,  
 RA Grohmann L.;  
 RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of  
 RT the respiratory chain from the inner mitochondrial membrane of  
 RT Solanum tuberosum.";  
 RL J. Biol. Chem. 269:2263-2269(1994).

CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
 CC TO BE UBIQUINONE.  
 CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.  
 CC -!- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER  
 CC MEMBRANE.  
 DR PIR: G49732; G49732.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1571 MW; 8C80B69EFA87AAC9 CRC64;  
 Query Match 20.8%; Score 25; DB 1; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
 QY 12 SNPTDD----DVSS 21  
 DB 2 SNATDETXLKDVSA 15  
 RESULT 7  
 ID TKNA\_RANRI STANDARD; PRT; 11 AA.  
 AC P29207;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RANAKININ (SUBSTANCE-P-RELATED PEPTIDE).  
 OS Rana ridibunda (Laughing frog) (Marsh frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.  
 OX NCBI\_TaxID=8406;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=92044543; PubMed=1658233;  
 RA O'Harte F., Burcher E., Lovas S., Smith D.D., Vaudry H., Conlon J.M.;  
 RT "Ranakinin: a novel NK1 tachykinin receptor agonist isolated with  
 RT neurokinin B from the brain of the frog Rana ridibunda.";  
 RL J. Neurochem. 57:2086-2091(1991).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1352 MW; 3A2460CC59D40B07 CRC64;  
 Query Match 20.0%; Score 24; DB 1; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 RTNPEDIY 10  
 DB 1 KPNPERFY 8  
 RESULT 8  
 ID TKND\_RANCA STANDARD; PRT; 11 AA.  
 AC P22691;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE RANATACHYKININ D (RTK D).  
OS Rana catesbeiana (Bull frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
OX NCBI\_TaxID=8400;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Intestine;  
RX MEDLINE=91254337; PubMed=2043143;  
RA Kozawa H., Hino J., Minamino N., Kangawa K., Matsuo H.;  
RT "Isolation of four novel tachykinins from frog (Rana catesbeiana)  
brain and intestine.";  
RL Biochem. Biophys. Res. Commun. 177:588-595(1991).  
RN [2]  
RP SEQUENCE.  
RC TISSUE=Intestine;  
RX MEDLINE=94023216; PubMed=8210506;  
RA Kangawa K., Kozawa H., Hino J., Minamino N., Matsuo H.;  
RT "Four novel tachykinins in frog (Rana catesbeiana) brain and  
intestine.";  
RL Regul. Pept. 46:81-88(1993).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR: JE0429; JE0429.  
DR PIR: D61033; D61033.  
DR InterPro: IPR002040; Tachykinin.  
DR PROSITE: PS00267; TACHYKININ; FALSE\_NEG.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 11 11 AMIDATION.  
SQ SEQUENCE 11 AA; 1350 MW; 3A34256C59D40B07 CRC64;

Query Match 20.0%; Score 24; DB 1; Length 11;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 RYNPEDIY 10  
DB 1 KPNPERFY 8

RESULT 9  
TRYP\_FELCA STANDARD; PRT; 16 AA.  
AC P81071;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE TRYPsin PRECURSOR (EC 3.4.21.4) (FRAGMENT).  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Pancreas;  
RX MEDLINE=97235546; PubMed=9080665;  
RA Steiner J.M., Medinger T.L., Williams D.A.;  
RT "Purification and partial characterization of feline trypsin.";  
RL Comp. Biochem. Physiol. 116B:87-93(1997).  
CC -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
TRYPsin FAMILY.  
DR InterPro: IPR001254; Trypsin.  
DR PROSITE: PS00240; TRYPsin\_DOM; PARTIAL.  
DR PROSITE: PS00134; TRYPsin\_HIS; PARTIAL.  
DR PROSITE: PS00135; TRYPsin\_SER; PARTIAL.  
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.  
FT PROPEP 1 8 ACTIVATION PEPTIDE.

FT CHAIN 9 >16 TRYPSIN.  
FT NON\_TER 16  
SQ SEQUENCE 16 AA; 1825 MW; A6D751BB58760A86 CRC64;

Query Match 20.0%; Score 24; DB 1; Length 16;  
Best Local Similarity 80.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 PTDDD 18  
DB 2 PIDDD 6

RESULT 10  
UPA2\_HUMAN STANDARD; PRT; 10 AA.  
AC P30088;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 10) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Plasma;  
RX MEDLINE=93092937; PubMed=1459097;  
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
RA Hochstrasser D.F.;  
RT "Plasma protein map: an update by microsequencing.";  
RL Electrophoresis 13:707-714(1992).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
PROTEIN IS: 4.4, ITS MW IS: 49 KDA.  
DR SWISS-2DPAGE; P30088; HUMAN.  
FT NON\_TER 1 1  
FT UNSURE 6 6  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1079 MW; 51AC54AAB7775B7 CRC64;

Query Match 19.2%; Score 23; DB 1; Length 10;  
Best Local Similarity 57.1%; Pred. No. 6.1e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 NPTDDVV 19  
DB 4 SPDDQV 10

RESULT 11  
FARI\_CALVO STANDARD; PRT; 12 AA.  
AC P41869;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CALLIMIRFAMIDE 1.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Thoracic ganglion;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2

RT neuropeptides (designated callifmrfamides) from the blowfly  
 RT Calliphora vomitoria.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR: E44787; E44787.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 12 12 AMIDATION.  
 SQ SEQUENCE 12 AA; 1390 MW; 2DDDB316914AB5A7 CRC64;

Query Match 19.2%; Score 23; DB 1; Length 12;  
 Best Local Similarity 42.9%; Pred. No. 7.5e+02;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 PSNPTDD 17  
 I: I:  
 Db 2 PNQPSDN 8

RESULT 12  
 TPL3\_PHYRO  
 ID TP13\_PHYRO STANDARD; PRT; 13 AA.  
 AC P04096;  
 DT 01-NOV-1986 (Rel. 03, Created)  
 DT 01-NOV-1986 (Rel. 03, Last sequence update)  
 DT 01-APR-1988 (Rel. 07, Last annotation update)  
 DE TRYPTOPHYLLIN-13.  
 OS Phylomedusa rohdei (Rohde's leaf frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
 OC Phylomedusa.  
 OX NCBI\_TaxID=8394;  
 RN [1]  
 RP SEQUENCE.  
 RA Montecucchi P.C., Gozzini L., Ersamer V.;  
 RT "Primary structure determination of a tryptophan-containing  
 RT tridecapeptide from Phylomedusa rohdei.";  
 RL Int. J. Pept. Protein Res. 27:175-182(1986).  
 DR PIR: A05174; A05174.  
 KW Amphibian skin.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 SQ SEQUENCE 13 AA; 1646 MW; 33BF3A212227773 CRC64;

Query Match 19.2%; Score 23; DB 1; Length 13;  
 Best Local Similarity 66.7%; Pred. No. 8.2e+02;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 PEDIYP 11  
 I: I:  
 Db 7 PPIYP 12

RESULT 13  
 UN02\_PINPS  
 ID UN02\_PINPS STANDARD; PRT; 13 AA.  
 AC P81667;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (N55) (FRAGMENTS).  
 OS Pinus pinaster (Maritime pine).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=71647;  
 RN [1]  
 RP SEQUENCE.  
 RA TISSUE=Needle;  
 RX MEDLINE=99274088; PubMed=10344291;  
 RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,  
 RA Frigerio J.-M., Plomion C.;  
 RT "Separation and characterization of needle and xylem maritime pine

RT proteins.";  
 RL Electrophoresis 20:1098-1108(1999).  
 CC -1- INDUCTION: BY WATER-STRESS.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 5.4, ITS MW IS: 43 KDA.  
 FT NON\_TER 1 1  
 FT NON\_CONS 7 8  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1559 MW; 96B5A43EF94B411 CRC64;

Query Match 19.2%; Score 23; DB 1; Length 13;  
 Best Local Similarity 62.5%; Pred. No. 8.2e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 SNPTDDV 19  
 I: I:  
 Db 3 SAPTRDNV 10

RESULT 14  
 UP51\_UPEIN  
 ID UP51\_UPEIN STANDARD; PRT; 13 AA.  
 AC P82036;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE UPERIN 5.1.  
 OS Uperoleia inundata (Floodplain toadlet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
 OC Uperoleia.  
 OX NCBI\_TaxID=104953;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RA TISSUE=Skin secretion;  
 RA Bradford A.M., Rafferty M.J., Bowie J.H., Tyler M.J., Wallace J.C.,  
 RA Adams G.W., Severini C.;  
 RT "Novel uperin peptides from the dorsal glands of the Australian  
 RT floodplain toadlet Uperoleia inundata.";  
 RL Aust. J. Chem. 49:475-484(1996).  
 CC -1- FUNCTION: UNKNOWN.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.  
 CC -1- MASS SPECTROMETRY: MW=1426; METHOD=FAB.  
 KW Amphibian skin.  
 SQ SEQUENCE 13 AA; 1457 MW; 290CA148C384B5A7 CRC64;

Query Match 19.2%; Score 23; DB 1; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 8.2e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPEDI 9  
 I: I:  
 Db 5 NPSDI 9

RESULT 15  
 CXAL\_CONAL  
 ID CXAL\_CONAL STANDARD; PRT; 16 AA.  
 AC P56639;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE ALPHA-CONOTOXIN AUIA.  
 OS Conus aulicus (Court cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=89437;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Venom;  
 RX MEDLINE=99003392; PubMed=9786965;

RA Luo S., Kulak J.M., Cartier G.E., Jacobsen R.B., Yoshikami D.,  
 RA Olivera B.M., McIntosh J.M.;  
 RT "Alpha-conotoxin AuIB selectively blocks alpha3 beta4 nicotinic  
 RT acetylcholine receptors and nicotine-evoked norepinephrine release."  
 RL J. Neurosci. 18:8571-8579(1998).  
 CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
 CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC  
 CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-4 SUBUNITS.  
 CC -|- MASS SPECTROMETRY: MW=1725.6; METHOD=ELECTROSPRAY.  
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
 KW Venom.  
 FT DISULFID 2 8  
 FT DISULFID 3 16  
 FT MOD.RES 16 16 AMIDATION  
 SQ SEQUENCE 16 AA; 1731 MW; 1E310FEB8FDC7001 CRC64;

Query Match 19.2%; Score 23; DB 1; Length 16;  
 Best Local Similarity 42.9%; Pred. No. 1e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CYRTNPE 7  
 Db 8 CFATNSD 14

Search completed: February 21, 2002, 16:52:20  
 Job time: 519 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: February 21, 2002, 16:43:01 ; Search time 60.03 seconds  
(without alignments)  
51.170 Million cell updates/sec

Title: US-08-753-851-5  
Perfect score: 120  
Sequence: 1 CYRTNPEDIYPSNPTDDVSS 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 5819

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL.17.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	26.7	13	5 Q917F8	Q917F8 drosophila
2	31.5	26.2	19	11 Q9QV20	Q9qv20 mus sp. per
3	30	25.0	21	4 Q16218	Q16218 homo sapien
4	28	23.3	20	2 Q9RLP2	Q9rlp2 neisseria g
5	28	23.3	21	6 Q18764	Q18764 sus scrofa
6	27	22.5	20	6 Q28079	Q28079 bos taurus
7	26	21.7	18	2 P82573	P82573 streptococc
8	26	21.7	18	9 Q38573	Q38573 bacterioph
9	26	21.7	18	11 P70650	P70650 mus musculu
10	26	21.7	19	8 Q9XNQ3	Q9xmq3 boophilus m
11	26	21.7	20	5 Q9TW18	Q9tw18 ceratilis c
12	26	21.7	20	5 Q9TW15	Q9tw15 ceratilis c
13	26	21.7	20	11 Q9QUX7	Q9qux7 rattus sp.
14	25	20.8	16	4 Q9UC52	Q9uc52 homo sapien
15	25	20.8	18	4 Q9UCL5	Q9ucl5 homo sapien
16	25	20.8	18	11 Q9QV11	Q9qv11 rattus sp.
17	25	20.8	18	11 Q9EQRL	Q9eqrl mus musculu
18	25	20.8	20	2 Q9R896	Q9r896 chlamydia t
19	25	20.8	21	3 Q9UTT6	Q9utt6 schizosacch

20	25	20.8	21	4 Q9UD48	Q9ud48 homo sapien
21	25	20.8	21	13 Q90609	Q90609 gallus gall
22	24.5	20.4	20	5 P82310	P82310 panulirus j
23	24.5	20.4	20	6 Q9TR87	Q9tr87 bos taurus
24	24	20.0	12	11 Q9CT94	Q9ct94 mus musculu
25	24	20.0	15	10 Q40562	Q40562 nicotiana t
26	24	20.0	15	10 Q40563	Q40563 nicotiana t
27	24	20.0	16	4 Q9NPQ7	Q9npq7 homo sapien
28	24	20.0	18	6 Q46865	Q46865 macropus ro
29	24	20.0	18	11 P70649	P70649 mus musculu
30	24	20.0	20	4 Q9UCK0	Q9uck0 homo sapien
31	24	20.0	20	10 Q9S8K0	Q9s8k0 solanum tub
32	24	20.0	21	11 Q9R1H1	Q9r1h1 rattus norv
33	23	19.2	12	5 P82619	P82619 periplaneta
34	23	19.2	13	8 Q9TKG6	Q9tkg6 lambia anta
35	23	19.2	15	1 Q9UWM1	Q9uwm1 methanospir
36	23	19.2	15	2 Q9R5D5	Q9r5d5 chromatium
37	23	19.2	16	2 Q9R5E9	Q9r5e9 haemophilus
38	23	19.2	16	6 Q9TQ29	Q9tq29 sus scrofa
39	23	19.2	17	4 Q9UWN3	Q9umn3 homo sapien
40	23	19.2	17	13 Q90737	Q90737 gallus gall
41	23	19.2	19	2 Q9R4J9	Q9r4j9 mycobacteri
42	23	19.2	19	6 Q9TRP4	Q9trp4 bos taurus
43	23	19.2	19	11 Q9Z1J5	Q9z1j5 rattus sp.
44	23	19.2	20	2 Q50180	Q50180 pseudomonas
45	23	19.2	20	3 P82262	P82262 aspergillus

#### ALIGNMENTS

RESULT 1  
Q917F8 ID Q917F8 PRELIMINARY; PRT; 13 AA.  
AC Q917F8;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CG18278 PROTEIN (FRAGMENT).  
GN CG18278.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadenot E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AF003818; AAG22274.1; -;  
 DR FlyBase: FBgn0033836; CG18278.  
 FT NON\_TER 1  
 SQ SEQUENCE 13 AA; 1554 MW; 5D52855A93735EB7 CRC64;

Query Match 26.7%; Score 32; DB 5; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYRTN 5  
 DB 7 CYRTN 11

RESULT 2  
 Q9QV20 PRELIMINARY; PRT; 19 AA.  
 AC Q9QV20;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE PEROXISOMAL MULTIFUNCTIONAL ENZYME (FRAGMENT).  
 OS Mus sp.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95040853; PubMed=7953067;  
 RA Stark A., Meijer J.;  
 RT "Purification and characterization of multifunctional enzyme from  
 RT mouse liver peroxisomes."  
 RL Comp. Biochem. Physiol. 108B:471-480(1994).  
 FT NON\_TER 1  
 FT NON\_TER 19  
 SQ SEQUENCE 19 AA; 2424 MW; 73E5F8EAD3A4EBD9 CRC64;

Query Match 26.2%; Score 31.5; DB 11; Length 19;  
 Best Local Similarity 53.3%; Pred. No. 3.7e+02;  
 Matches 8; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 YRTNPEDIYPSNPTD 16  
 DB 2 YRQNP-DIQLPSPD 15

RESULT 3  
 Q16218 PRELIMINARY; PRT; 21 AA.  
 AC Q16218;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE DELETED IN COLORECTAL CANCER PROTEIN (FRAGMENT).  
 GN DCC.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94320088; PubMed=8044801;  
 RA Reale M.A., Hu G., Zafar A.I., Getzenberg R.H., Levine S.M.,  
 RA Fearon E.R.;  
 RT "Expression and alternative splicing of the deleted in colorectal  
 RT cancer (DCC) gene in normal and malignant tissues.";  
 RL Cancer Res. 54:4493-4501(1994).  
 DR EMBL: S73204; AAB31484.2; -;  
 FT NON\_TER 1  
 FT NON\_TER 21  
 SQ SEQUENCE 21 AA; 2337 MW; D2C3578D41548440 CRC64;

Query Match 25.0%; Score 30; DB 4; Length 21;  
 Best Local Similarity 62.5%; Pred. No. 6.8e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 TNPEDIYP 11  
 DB 2 TDPVDIYP 9

RESULT 4  
 Q9RLP2 PRELIMINARY; PRT; 20 AA.  
 AC Q9RLP2;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE PUTATIVE CYSTEINEYL T-RNA SYNTHETASE (FRAGMENT).  
 OS Neisseria gonorrhoeae.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=485;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WR302;  
 RA Stein D.C., Gunn J.S.;  
 RT "Use of a non-selective transformaton technique to construct a  
 RT multiple restriction-modification deficient mutant of *Neisseria*  
 RT gonorrhoeae.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U42459; AAD56589.1; -;  
 FT NON\_TER 20  
 FT NON\_TER 20  
 SQ SEQUENCE 20 AA; 2245 MW; 4C893C704879FFD3 CRC64;

Query Match 23.3%; Score 28; DB 2; Length 20;  
 Best Local Similarity 42.9%; Pred. No. 1.3e+03;  
 Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 YRTNPEDIYPSNPT 15  
 DB 5 YNTLPKPEPSPPS 18

RESULT 5  
 O18764 PRELIMINARY; PRT; 21 AA.  
 AC O18764;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR, BETA 4 SUBUNIT (FRAGMENT).  
 GN CHRNA4.

OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=98161940; PubMed=9501320;
RA Tammien I.;
RT "Genetic mapping of CHNA3 and CHRNA4 to pig chromosome 7 extends the
RT syntenic conservation with human chromosome 15 and mouse chromosome
RT 9.";
RL Mamm. Genome 9:263-264(1998).
DR EMBL; AF007797; AAC13376.1; -.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2196 MW; EA98A064013C0FFD CRC64;

Query Match 23.3%; Score 28; DB 6; Length 21;
Best Local Similarity 38.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 YRTNPEDIYPSNP 14
Db 6 YFVNPASAAPKSP 18

RESULT 6
Q28079 PRELIMINARY; PRT; 20 AA.
AC Q28079;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE EPITHELIAL MUCIN (FRAGMENT).
GN MUC1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LACTATING MAMMARY GLAND;
RX MEDLINE=96351712; PubMed=8747930;
RA Spicer A.P., Duhig T., Chilton B.S., Gendler S.J.;
RT "Analysis of mammalian MUC1 genes reveals potential functionally
RT important domains.";
RL Mamm. Genome 6:885-888(1995).
DR EMBL; L41553; AAB48100.1; -.
FT NON_TER 1
FT NON_TER 20
SQ SEQUENCE 20 AA; 2055 MW; 0E14465183AB743B CRC64;

Query Match 22.5%; Score 27; DB 6; Length 20;
Best Local Similarity 42.9%; Pred. No. 1.8e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 TNPEDIYPSNPSTD 17
Db 4 TSPNRKTSAPADD 17

RESULT 7
P82573 PRELIMINARY; PRT; 18 AA.
AC P82573;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE URIDINE KINASE (EC 2.7.1.48) (URIDINE MONOPHOSPHOKINASE) (FRAGMENTS).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC STRAIN-JRS4;

```

```

RA Hogan D.A., Du P., Stevenson T.I., Whittom M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
RT proteins.";
RL Submitted (MAY-2000) to the SWISS-PROT data bank.
CC -|- CATALYTIC ACTIVITY: ATP + URIDINE = ADP + UMP.
CC -|- CATALYTIC ACTIVITY: ATP + CYTIDINE = ADP + CMP.
CC -|- PATHWAY: PYRIMIDINE SALVAGE PATHWAY.
CC -|- MASS SPECTROMETRY: MW=24334.01; METHOD=ELECTROSPRAY.
CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
KW Transferase; Kinase.
FT NON_TER 1
FT NON_CONS 10
FT NON_TER 18
FT NON_TER 18
SQ SEQUENCE 18 AA; 1991 MW; BA532B372051EDC6 CRC64;

Query Match 21.7%; Score 26; DB 2; Length 18;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 TDDDV 19
Db 5 TDDDI 9

RESULT 8
Q38573 PRELIMINARY; PRT; 18 AA.
ID Q38573;
AC Q38573;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE COAT GENE PROTEIN (FRAGMENT).
GN COAT GENE.
OS Bacteriophage Kul.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
OC Levivirus.
OX NCBI_TaxID=12021;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96207403; PubMed=8615017;
RA Groeneveld H., Oudot F., van Duin J.V.;
RT "RNA phage KUI has an insertion of 18 nucleotides in the start codon
RT of its lysis gene.";
RL Virology 218:141-147(1996).
DR EMBL; S81763; AAD14371.1; -.
DR HSSP; P07234; IUNA.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 18 AA; 1825 MW; 2D4092DC226D8904 CRC64;

Query Match 21.7%; Score 26; DB 9; Length 18;
Best Local Similarity 55.6%; Pred. No. 2.2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 13 NPTDDDVSS 21
Db 4 NPVADAISS 12

RESULT 9
P70650 PRELIMINARY; PRT; 18 AA.
ID P70650;
AC P70650;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE SYNTAXIN 3 ISOFORM D-2 (FRAGMENT).
GN SYN3D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FOREBRAIN;  
 RX MEDLINE=95321968; PubMed=7598732;  
 RA Ibaraki K., Horikawa H.P., Morita T., Mori H., Sakimura K.,  
 RA Mishina M., Saisu H., Abe T.;  
 RT "Identification of four different forms of syntaxin 3";  
 RL Biochem. Biophys. Res. Commun. 211:997-1005(1995).  
 DR EMBL; S78677; AAC17462.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 18 AA; 2043 MW; 289FBCA184495155 CRC64;

Query Match 21.7%; Score 26; DB 11; Length 18;  
 Best Local Similarity 55.6%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 PEDIYSPNP 14  
 || | | |  
 DB 1 PEPKLPNP 9

RESULT 10

Q9XNQ3 ID Q9XNQ3 PRELIMINARY; PRT; 19 AA.  
 AC Q9XNQ3;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1) (FRAGMENT).  
 GN COI.  
 OS Boophilus microplus (Cattle tick).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
 OX NCBI\_TaxID=6941;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N;  
 RX MEDLINE=99297341; PubMed=10368952;  
 RA Campbell N.J.H., Barker S.C.;  
 RT "The novel mitochondrial gene arrangement of the cattle tick,  
 RT Boophilus microplus: fivefold tandem repetition of a coding region";  
 RL Mol. Biol. Evol. 16:732-740(1999).  
 CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY  
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. COI IS THE  
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN  
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2  
 CC AND COPPER B (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) - 2 H(2)O + 4  
 CC FERRICYTOCHROME C.  
 CC -!- COFACTOR: TWO HEME (A AND A3) GROUPS AND COPPER B (BY SIMILARITY).  
 CC -!- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
 DR EMBL; AF110613; AAD28396.1; -;  
 DR InterPro; IPR000883; COX1.  
 DR Pfam; PF00115; COX1; 1.  
 KW Copper; Heme; Inner membrane; Mitochondrion; Oxidoreductase;  
 KW Respiratory chain; Transmembrane.  
 FT NON\_TER 19  
 SQ SEQUENCE 19 AA; 2249 MW; C58C9C41FCAB0FED CRC64;

Query Match 21.7%; Score 26; DB 8; Length 19;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+03;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRTNPEDI 9  
 || | | |

Db 2 YSTNHKDI 9

RESULT 11

Q9TWI8 ID Q9TWI8 PRELIMINARY; PRT; 20 AA.  
 AC Q9TWI8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE MAJOR CUTICULAR PROTEIN (FRAGMENT).  
 OS Ceratitis capitata (Mediterranean fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Tephritoidea; Tephritidae; Ceratitis.  
 OX NCBI\_TaxID=7213;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95279142; PubMed=7759285;  
 RA Tsiliras A., Koromantzou E., Allen G., Dimitriadis G.;  
 RT "Partial N-terminal sequences of larval cuticular proteins from the  
 RT dipteran Ceratitis capitata";  
 RL Hereditas 122:79-83(1995).  
 SQ SEQUENCE 20 AA; 2408 MW; 2BC25E2F9EC15A80 CRC64;

Query Match 21.7%; Score 26; DB 5; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPEDIY 10  
 || | | |  
 DB 14 NPENFY 19

RESULT 12

Q9TWI5 ID Q9TWI5 PRELIMINARY; PRT; 20 AA.  
 AC Q9TWI5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE MAJOR CUTICULAR PROTEIN (FRAGMENT).  
 OS Ceratitis capitata (Mediterranean fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Tephritoidea; Tephritidae; Ceratitis.  
 OX NCBI\_TaxID=7213;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95279142; PubMed=7759285;  
 RA Tsiliras A., Koromantzou E., Allen G., Dimitriadis G.;  
 RT "Partial N-terminal sequences of larval cuticular proteins from the  
 RT dipteran Ceratitis capitata";  
 RL Hereditas 122:79-83(1995).  
 SQ SEQUENCE 20 AA; 2329 MW; C6ABB0B44ED70B88 CRC64;

Query Match 21.7%; Score 26; DB 5; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPEDIY 10  
 || | | |  
 DB 14 NPDSIY 19

RESULT 13

Q9QUX7 ID Q9QUX7 PRELIMINARY; PRT; 20 AA.  
 AC Q9QUX7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE CARBOXYLESTERASE ISOZYME (FRAGMENT).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96170127; PubMed=8597091;  
RA Satoh T., Hosokawa M.;  
RT "Molecular aspects of carboxylesterase isoforms in comparison with  
RT other esterases."  
RL Toxicol. Lett. 82:439-445(1995).  
SQ SEQUENCE 20 AA; 2133 MW; 435160FFA80E086D CRC64;

Query Match 21.7%; Score 26; DB 11; Length 20;  
Best Local Similarity 80.0%; Pred. No. 2.5e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 YPSNP 14  
|||:  
Db 1 YPSSP 5

RESULT 14  
Q9UC52  
ID Q9UC52 PRELIMINARY; PRT; 16 AA.  
AC Q9UC52;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE CARBOXYLESTERASE ISOZYME (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96170127; PubMed=8597091;  
RA Satoh T., Hosokawa M.;  
RT "Molecular aspects of carboxylesterase isoforms in comparison with  
RT other esterases."  
RL Toxicol. Lett. 82:439-445(1995).  
SQ SEQUENCE 16 AA; 1599 MW; 5DA6286D3C6BFFED CRC64;

Query Match 20.8%; Score 25; DB 4; Length 16;  
Best Local Similarity 62.5%; Pred. No. 2.8e+03;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 PSNPTDDD 18  
|||:  
Db 3 PSPVVDD 10

RESULT 15  
Q9UC15  
ID Q9UC15 PRELIMINARY; PRT; 18 AA.  
AC Q9UC15;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 22 KDA FORM (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93091816; PubMed=1726837;  
RA Roghani M., Segovia B., Whitechurch O., Binoux M.;  
RT "Purification from human cerebrospinal fluid of insulin-like growth

RT factor binding proteins (IGFBPs). Isolation of IGFBP-2, an altered  
RT form of IGFBP-3 and a new IGFBP species."  
RL Growth Regul. 1:123-130(1991).  
SQ SEQUENCE 18 AA; 1948 MW; AC3F4170E95081E3 CRC64;

Query Match 20.8%; Score 25; DB 4; Length 18;  
Best Local Similarity 26.7%; Pred. No. 3.1e+03;  
Matches 4; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 7 EDIYPSNPTDDDDVSS 21  
: | | |:  
Db 1 DSFVXPXEPSDEKALS 15

Search completed: February 21, 2002, 16:51:54  
Job time: 533 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:39:06 ; Search time 64.88 Seconds  
(without alignments)  
23.976 Million cell updates/sec

Title: US-08-753-851-5  
Perfect score: 120  
Sequence: 1 CYRTNPEDIYPSNPTDDVSS 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 206698

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

```
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	21	15 AAR53477	CD44 peptide CD44-
2	36.5	30.4	21	19 AAW42092	BS106 synthetic po
3	36.5	30.4	21	20 AAY13470	BS106 polypeptide
4	36	30.0	16	18 AAW14606	Human antithrombin
5	35.5	29.6	16	20 AAY43350	GTF antigenic epit
6	33	27.5	7	19 AAW54457	Human native angio
7	33	27.5	8	19 AAW54458	Human native angio
8	33	27.5	10	21 AAB06281	proSKI-1 putative
9	33	27.5	14	21 AAB08529	Amino acid sequenc
10	33	27.5	18	20 AAY05230	Human Stat6 antige
11	33	27.5	20	22 AAG64195	Protein sequence 2

12	33	27.5	20	22 AAB99270	Peptide sequence 2
13	32	26.7	11	18 AAW34298	Kringle 5 peptide
14	32	26.7	11	21 AAB01899	Human plasminogen
15	32	26.7	11	22 AAB92094	Laminin fragment S
16	32	26.7	11	22 AAB36569	Mammalian kringle
17	32	26.7	12	22 AAB92089	Laminin fragment S
18	32	26.7	12	22 AAB92093	Laminin fragment S
19	32	26.7	12	22 AAB36564	Mammalian kringle
20	32	26.7	12	22 AAB36568	Mammalian kringle
21	32	26.7	16	11 AAY17924	Synthetic peptide
22	32	26.7	16	11 AAR07527	B-cell epitope pe
23	32	26.7	16	18 AAW09992	HTLV-1 envelope pr
24	32	26.7	16	19 AAW37101	HTLV-1 envelope an
25	32	26.7	16	20 AAY00798	HTLV1 envelope pep
26	32	26.7	16	21 AAY94586	Two tandem copies
27	32	26.7	19	18 AAW31054	Mugwort pollen all
28	31	25.8	10	18 AAW43376	Immunogenic HLA-A2
29	31	25.8	10	21 AAY53097	Anticoagulant pept
30	31	25.8	10	22 AAM09237	HLA-A1 decamer #10
31	31	25.8	10	22 AAM09297	HLA-B *0702 decame
32	31	25.8	10	22 AAM10474	HLA-A *0201 decame
33	31	25.8	10	22 AAM13407	HLA-A26 decamer #4
34	31	25.8	12	21 AAY53086	Anticoagulant pept
35	31	25.8	13	21 AAB10012	H. pylori beta-ure
36	31	25.8	13	22 AAB86060	H. pylori beta-ure
37	31	25.8	13	22 AAB86092	H. pylori beta-ure
38	31	25.8	15	20 AAW89494	Conopeptide AuiB.
39	31	25.8	15	22 AAG88514	HER2/NEU DR superm
40	31	25.8	17	20 AAW89159	Anti-p53 monoclonal
41	31	25.8	19	14 AAF54356	RAE 40.1. Ambrosi
42	31	25.8	20	19 AAW56585	Synthetic TcaAII p
43	31	25.8	20	22 AAB55809	PL peptide #20. S
44	31	25.8	20	22 AAB57671	DNAM-1 PL peptide.
45	31	25.8	20	22 AAB58081	DNAM-1 PL peptide.

#### ALIGNMENTS

RESULT 1  
AAR53477  
ID AAR53477 standard; peptide; 21 AA.  
XX AAR53477;  
XX AC  
XX  
DT 01-DEC-1994 (first entry)  
XX  
DE CD44 peptide CD44-5.  
XX  
KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autoimmune disease;  
KW immunosuppressive.  
XX  
OS Homo sapiens.  
XX  
PN WO9409811-A.  
XX  
PD 11-MAY-1994.  
XX  
PF 29-OCT-1993; 93WO-US10412.  
XX  
PR 30-OCT-1992; 92US-0973339.  
XX  
(UYDU-) UNIV DUKE.  
XX  
PA Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
XX  
PI GTF antigenic epit  
XX  
DR Human native angio  
XX  
XX WPI; 1994-167121/20.  
XX  
PT Use of CD44 protein and new peptide.derivs - for developing prods  
XX for inflammation, immune-mediated tissue damage and tumour cell  
XX metastasis

PS Claim 4; Page 14; 83pp; English.

CC The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 161-180 of the CD44 protein sequence.

XX Sequence 21 AA;

Query Match 100.0%; Score 120; DB 15; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-11;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CYRTNPEDIYPSNPTDDVSS 21  
Db 1 cyrtnpediypsnptddvss 21

RESULT 2

AAW42092 ID AAW42092 standard; Peptide: 21 AA.

XX AAW42092;

DT 25-SEP-1998 (first entry)

DE BS106 synthetic polypeptide 4.

XX Human; BS106; breast tissue gene; breast cancer; detection marker.

OS Homo sapiens.

XX WO9818945-A1.

PN 07-MAY-1998.

XX 31-OCT-1997; 97WO-US19836.

PR 31-OCT-1996; 96US-0742067.

PA (ABBO ) ABBOTT LAB.

PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
PI Roberts-Rapp L, Russell JC, Stroupe SD;

DR WPI; 1998-272234/24.

PT Detecting BS106 gene products - useful for developing products for detecting, staging, preventing, treating or determining predisposition to e.g. breast cancer

PS Claim 17; Page 93; 114pp; English.

XX BS106 is a breast tissue gene with which breast cancer and related diseases are associated. The BS106 polypeptides AAW42088-W42092 can be utilized in a variety of assays for the detection of antibodies to breast tissue. They can also be used as immunogens for the detection of antibodies to breast tissue. The products and methods discussed in this invention can be used for detecting, diagnosing, staging, preventing or treating, or determining predisposition to diseases or conditions of the breast such as breast cancer.

XX Sequence 21 AA;

Query Match 30.4%; Score 36.5; DB 19; Length 21;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 7; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 6 PEDIYP-SNPTDDD 18  
Db 3 padtypatgpadde 16

RESULT 3

AAW13470 ID AAW13470 standard; Protein: 21 AA.

XX AAW13470;

DT 26-JUL-1999 (first entry)

DE BS106 polypeptide sequence.

XX BS106; breast; marker; detection; diagnosis; breast cancer.

OS Homo sapiens.

XX WO9923230-A1.

PN 14-MAY-1999.

XX 19-OCT-1998; 98WO-US22020.

PR 31-OCT-1997; 97US-0962094.

PA (ABBO ) ABBOTT LAB.

PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
PI Roberts-Rapp L, Russell JC, Stroupe SD;

DR WPI; 1999-326991/27.

XX New isolated BS106 nucleic acids

PT Claim 17; Page 93; 105pp; English.

XX The invention provides BS106 polynucleotides (AAW55578-82) and polypeptides (AAW13466-470). The BS106 nucleic acids correspond to a breast tissue gene and can be used as markers for breast tissue disease. The products can be used for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating, or determining diseases or conditions associated with BS106, especially breast cancer.

XX Sequence 21 AA;

Query Match 30.4%; Score 36.5; DB 20; Length 21;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 7; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 6 PEDIYP-SNPTDDD 18  
Db 3 padtypatgpadde 16

RESULT 4

AAW14606

ID AAW14606 standard; peptide: 16 AA.

XX AAW14606;

DT 09-DEC-1997 (first entry)

DE Human antithrombin III carboxy-terminal extension.

XX Human; antithrombin III; mutant; carboxy-terminal; extension;  
KW sulphated tyrosine; prevention; treatment; thrombosis; disease.  
XX Synthetic.



FH Key: Location/Qualifiers  
 FT Region 1..3  
 FT /note= "carboxy-terminus of wild type human  
 FT antithrombin III"  
 XX  
 XX JP09071600-A.  
 PN 18-MAR-1997.  
 PD  
 PD 06-SEP-1995; 95JP-0228694.  
 XX  
 XX 06-SEP-1995; 95JP-0228694.  
 PR  
 XX (EISA ) EISAI CO LTD.  
 PA  
 XX WPI; 1997-231188/21.  
 DR  
 XX Human antithrombin III mutant with C-terminal peptide containing a  
 XX sulphated tyrosine - useful for treatment and prevention of thrombic  
 PT disease  
 PT  
 XX Disclosure; Page 6; lpp; Japanese.  
 PS  
 CC A human antithrombin III (AF III) mutant, comprising a  
 CC carboxy-terminal sulphated tyrosine peptide extension, i.e. the  
 CC present peptide, can be used to prevent and/or treat thrombotic  
 CC diseases.  
 CC  
 XX Sequence 16 AA;  
 SQ

Query Match 30.0%; Score 36; DB 18; Length 16;  
 Best Local Similarity 37.5%; Pred. NO. 53;  
 Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 CYRTNPEDIYPSNPTD 16  
 | : | : | : | :  
 Db 1 cvktdlyyyeedte 16

RESULT 5  
 AAY43350  
 ID AAY43350 standard; peptide; 16 AA.  
 XX  
 AC AAY43350;  
 XX  
 DT 26-JAN-2000 (first entry)  
 XX  
 DE GTF antigenic epitope.  
 XX  
 KW GTF; antigenic epitope; glucosyltransferase; immunogenic epitope; teeth;  
 KW immunogenic composition; streptococcal glucan; antibody response;  
 KW cariogenic bacteria; mutans streptococci; colonisation; caries;  
 KW dental caries; immunisation; therapy.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9952548-A2.  
 XX  
 PD 21-OCT-1999.  
 XX  
 PF 09-APR-1999; 99WO-US07828.  
 XX  
 PR 10-APR-1998; 98US-0081315.  
 XX  
 PA (LEES/) LEES A.  
 PA (TAUB/) TAUBMAN M A.  
 PA (SMIT/) SMITH D J.  
 XX  
 PI Lees A, Taubman MA, Smith DJ;  
 XX  
 DR WPI; 1999-620289/53.  
 XX

PT Immunogenic compositions for control of dental caries, based on  
 PT Streptococcus mutans components, particularly for vaccination of  
 PT infants  
 XX  
 XX Claim 11; Page 44; 54pp; English.  
 PS  
 CC This sequence represents a glucosyltransferase (GTF) immunogenic epitope,  
 CC and can be used in the immunogenic composition of the invention. The  
 CC immunogenic composition (A) comprises at least one streptococcal glucan  
 CC (SG) and at least one component (I) covalently coupled, (in)directly, to  
 CC SG. (A) elicits antibodies (Ab) to both SG and (I). (A), and related  
 CC compositions, are used to induce an antibody response to cariogenic  
 CC bacteria, collectively called "mutans streptococci" (i.e. any of  
 CC Streptococcus mutans, S. rattus, S. cricetus or S. sobrinus), especially  
 CC to prevent colonisation, to reduce numbers of bacteria or to reduce  
 CC incidence of caries, in humans or other animals. The composition can  
 CC therefore be used for the treatment or prevention of caries (particularly  
 CC dental caries). Abs raised against (A) can be used: (a) to detect SG, and  
 CC (b) for passive immunisation, e.g. applied directly to the surface of  
 CC teeth. SG is a poorly immunogenic type 2 T-cell independent antigen but  
 CC when coupled to (I) it produces a conjugate which induces significant  
 CC immunity to both components, resulting in immunological memory and  
 CC long-lasting antibody production against SG, in adults and children.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 29.6%; Score 35.5; DB 20; Length 16;  
 Best Local Similarity 58.8%; Pred. No. 63;  
 Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 2;  
 QY 6 PEDIYPSNPTD-DDVSS 21  
 | | | | : | | | | |  
 Db 2 ped--ptdptdpgdps 16

RESULT 6  
 AAW54457  
 ID AAW54457 standard; protein; 7 AA.  
 XX  
 AC AAW54457;  
 XX  
 DT 28-AUG-1998 (first entry)  
 XX  
 DE Human native angiostatin C-terminal sequence 1.  
 XX  
 KW Angiostatin; plasminogen; human; thiol donor; activator; tumour;  
 KW angiogenic disease; rheumatoid arthritis; atherosclerosis; eye disorder;  
 KW cardiovascular; cerebral vascular; diabetes; immune disease; plasmin;  
 KW chronic inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9815574-A1.  
 XX  
 PD 16-APR-1998.  
 XX  
 PF 17-SEP-1997; 97WO-US16539.  
 XX  
 PR 17-SEP-1996; 96US-0710305.  
 XX  
 PA (NOUN ) UNIV NORTHWESTERN.  
 XX  
 PI Gately ST, Soff G, Twardowski P;  
 XX  
 DR WPI; 1998-240779/21.  
 XX  
 XX Generating angiostatin in vitro or in vivo - by treating plasminogen  
 PT with activator and thiol donor, for treatment of angiogenic disease,  
 PT e.g. tumours, arthritis, eye disease etc.  
 XX  
 XX Claim 42; Page 79; 98pp; English.  
 PS

CC This represents the human native angiotensin C-terminal sequence. This  
 CC was deduced to be a plasminogen fragment. The invention provides a  
 CC method for in vitro generation of angiotensin which comprises treating  
 CC plasminogen with a plasminogen activator and a sulphhydryl (thiol)  
 CC donor. The thiol donor converts the plasmin to angiotensin which can be  
 CC used for the treatment of angiogenic disease. The generated  
 CC angiotensin is a protein that is a fragment of plasminogen having the  
 CC same N-terminal amino acid sequence as plasmin and a C-terminus in  
 CC kringle 5. The generated angiotensins are used to treat angiogenic  
 CC disease, e.g. tumours and their metastases, rheumatoid arthritis,  
 CC atherosclerosis, eye, cardiovascular, cerebral vascular or diabetes  
 CC associated disorders, or immune diseases such as chronic inflammation.  
 CC A DNA that encodes the protein optionally linked to expression control  
 CC elements can be used in gene therapy to treat the same diseases.  
 CC Antibodies that bind selectively to the native angiotensin are used in  
 CC standard immunoassays for detection or determination of angiotensin for  
 CC diagnosis or monitoring of angiogenesis, and affinity purification. These  
 CC methods allow angiotensin to be produced on a large scale.

XX Sequence 7 AA;

Query Match 27.5%; Score 33; DB 19; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 4.3e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CYRTNP 6  
 ||||  
 Db 1 cytnp 6

# RESULT 7

AAW54458  
 ID AAW54458 standard; protein; 8 AA.

XX AAW54458;

XX 28-AUG-1998 (first entry)

XX Human native angiotensin C-terminal sequence 2.

XX Angiotensin; plasminogen; human; thiol donor; activator; tumour;  
 KW angiogenic disease; rheumatoid arthritis; atherosclerosis; eye disorder;  
 KW cardiovascular; cerebral vascular; diabetes; immune disease; plasmin;  
 KW chronic inflammation.

OS Homo sapiens.

XX WO9815574-A1.

XX 16-APR-1998.

XX 17-SEP-1997; 97WO-US16539.

XX 17-SEP-1996; 96US-0710305.

XX (NOUN ) UNIV NORTHWESTERN.

XX Gately ST, Soff G, Twardowski P;

XX WPI; 1998-240779/21.

XX Generating angiotensin in vitro or in vivo - by treating plasminogen  
 PT with activator and thiol donor, for treatment of angiogenic disease,  
 PT e.g. tumours, arthritis, eye disease etc.

XX Claim 42; Page 79; 98pp; English.

XX This represents the human native angiotensin C-terminal sequence. This  
 CC was deduced to be a plasminogen fragment. The invention provides a  
 CC method for in vitro generation of angiotensin which comprises treating  
 CC plasminogen with a plasminogen activator and a sulphhydryl (thiol)  
 CC donor. The thiol donor converts the plasmin to angiotensin which can be

CC used to for the treatment of angiogenic disease. The generated  
 CC angiotensin is a protein that is a fragment of plasminogen having the  
 CC same N-terminal amino acid sequence as plasmin and a C-terminus in  
 CC kringle 5. The generated angiotensins are used to treat angiogenic  
 CC disease, e.g. tumours and their metastases, rheumatoid arthritis,  
 CC atherosclerosis, eye, cardiovascular, cerebral vascular or diabetes  
 CC associated disorders, or immune diseases such as chronic inflammation.  
 CC A DNA that encodes the protein optionally linked to expression control  
 CC elements can be used in gene therapy to treat the same diseases.  
 CC Antibodies that bind selectively to the native angiotensin are used in  
 CC standard immunoassays for detection or determination of angiotensin for  
 CC diagnosis or monitoring of angiogenesis, and affinity purification. These  
 CC methods allow angiotensin to be produced on a large scale.

XX Sequence 8 AA;

Query Match 27.5%; Score 33; DB 19; Length 8;  
 Best Local Similarity 83.3%; Pred. No. 4.3e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CYRTNP 6  
 ||||  
 Db 1 cytnp 6

# RESULT 8

AAW06281  
 ID AAW06281 standard; peptide; 10 AA.

XX AAW06281;

XX 03-OCT-2000 (first entry)

XX proSKI-1 putative zymogen cleavage site.

XX Human; pro-brain-derived neurotrophic factor; proBDNF;  
 KW subtilisin-kexin isoenzyme 1; SKI-1; proSKI-1 zymogen processing;  
 KW antilipase; cytotaxin; vasotropic; SKI-1 inhibitor;  
 KW hypercholesterolemia; liver steatosis; Ras-dependent cancer;  
 KW restenosis; amyloid protein formation.

XX Homo sapiens.

OS Mus sp.

OS Rattus sp.

XX Key Location/Qualifiers

XX Cleavage-site 6..7

XX WO200026348-A2.

XX 11-MAY-2000.

XX 04-NOV-1999; 99WO-CA01058.

XX 04-NOV-1998; 98CA-2249648.

XX (RECL-) INST RECH CLINIQUES MONTREAL.

XX Seidah N, Chretien M, Marcinkiewicz M, Laaksonen R, Davignon J;

XX WPI; 2000-365601/31.

XX Novel soluble proteic fragment of subtilisin-kexin isoenzyme for  
 PT producing a polypeptide useful for treating hypercholesterolemia, liver  
 PT steatosis and amyloidosis, comprises a specific amino acid sequence -

XX Example 1; Page 25; 119pp; English.

XX The present sequence is a possible cleavage site in proSKI-1  
 CC which may be targeted by a zymogen to form mature subtilisin-kexin  
 CC isoenzyme 1 (SKI-1). The conserved site was identified by alignment of  
 CC SKI-1 with other subtilases. SKI-1 is a type-1 membrane-bound proteinase

CC which cleaves pro-brain-derived neurotrophic factor. Peptides which bind  
 CC to and are cleaved by SKI-1 may be used for monitoring SKI-1 activity,  
 CC for screening inhibitors of SKI-1 activity, or for screening enhancers of  
 CC SKI-1 activity. Protein fragments of SKI-1 which bind to the SKI-1  
 CC catalytic site may be used as inhibitors of SKI-1 activity. They may be  
 CC used to treat diseases involving overexpression of SKI-1 or SKI-1  
 CC substrate. Such diseases include hypercholesterolemia, high levels of  
 CC fatty acids, lipids or farnesyl pyrophosphate, liver steatosis,  
 CC Ras-dependent cancer, restenosis and amyloid protein formation.  
 XX  
 SQ Sequence 10 AA;

Query Match 27.5%; Score 33; DB 21; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 87;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3: RTNPEDIYPS 12  
 | | | | |  
 Db 1 rmpssdyps 10

RESULT 9  
 AAB08529  
 ID AAB08529 standard; Peptide; 14 AA.

XX AC AAB08529;

XX 20-DEC-2000 (first entry)

XX Amino acid sequence of a binding partner peptide of p47-phox.

DE Conformational state; post-translational modification; enzyme activity;  
 KW p47-phox; protein kinase C.  
 KW Unidentified.

XX WO200050901-A1.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-GB00668.

XX 25-FEB-1999; 99GB-0004395.

XX (FLUO-) FLUORESCENCE LTD.

XX Colyer J, Craig RK;

XX WPI; 2000-565475/52.

XX Determining the conformational state of a protein, comprises contacting  
 PT the protein with a labeled binding protein and assessing the labeling  
 PT of the protein -

XX Example 2; Page 30; 56pp; English.

XX The specification describes a method for determining the conformational  
 CC state of a protein. The method uses at least one labelled binding partner  
 CC capable of binding to the protein in a manner dependent on the  
 CC conformational state of the protein. The method is for detecting the  
 CC conformational state of a protein, for detecting post-translational  
 CC modifications of proteins, and for determining the activity of an enzyme.  
 CC The present sequence represents a binding partner of human p47-phox.  
 CC p47-phox is a substrate of protein kinase C. Detection of the  
 CC conformational changes in p47-phox is used to measure protein kinase C  
 CC activity, using the method of the invention.  
 XX

SQ Sequence 14 AA;

Query Match 27.5%; Score 33; DB 21; Length 14;  
 Best Local Similarity 35.7%; Pred. No. 1.3e+02;

Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 Qy 1 CYRTNPEDIYPSNP 14  
 | | | | |  
 Db 1 cqrskpqavpprp 14

RESULT 10  
 AAY05230  
 ID AAY05230 standard; peptide; 18 AA.

XX AC AAY05230;

XX 18-JUN-1999 (first entry)

XX Human Stat6 antigenic fragment.

DE Stat6; Stat6b; human; signal transducers and activators of transcription;  
 KW isoform; myeloid cancer; asthma; sarcoma; scleroderma; fibrotic disease;  
 KW bone marrow fibrosis; AIDS; Stat6c

XX Homo sapiens.

XX WO9910493-A1.

XX 04-MAR-1999.

XX 27-AUG-1998; 98WO-US17821.

XX 05-JAN-1998; 98US-0070397.

XX 27-AUG-1997; 97US-0056075.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Larocheville WJ, Patel B, Pierce JH;

XX WPI; 1999-214517/18.

XX New isoforms of Stat6 - having differential effects on the  
 PT modulation of Stat6 activity in cells

XX Example 1; Page 20; 88pp; English.

XX This sequence is an antigenic fragment of an isoform of  
 CC human Stat6 (signal transducers and activators of transcription), of the  
 CC invention. The detection and/or quantitation of DNA or  
 CC mRNA encoding Stat6 and/or Stat6b and/or Stat6c can be used to detect  
 CC differential expression of Stat6 isoforms in numerous diseases, including  
 CC myeloid cancer, asthma, sarcoma, scleroderma, bone marrow fibrosis,  
 CC fibrotic diseases and AIDS. The nucleic acids can be used to screen  
 CC genomic or cDNA libraries or to identify complementary sequences. The  
 CC identification of the genetic locus of the Stat6 gene can be used for  
 CC detection of chromosomal aberrations and translocations involving the  
 CC Stat6 gene. Antibodies against the isoforms can be used to detect the  
 CC presence of Stat6 and/or Stat6b and/or Stat6c in a sample. Because of the  
 CC variation of the roles of Stat6b and Stat6c in regulating gene  
 CC transcription the isolated and purified forms can be used to study gene  
 CC regulation and in screening assays for identifying drug candidates which  
 CC may be used as agonists or antagonists. The two polypeptides may also be  
 CC used in gene therapy protocols. In particular, Stat6b and/or Stat6c can  
 CC therapeutically modulate the development and differentiation of B and  
 CC T cells and can enhance IL-4 immunological function in immunocompromised  
 CC individuals. Stat6 activation correlates with functional responses  
 CC induced by interleukin-4 (IL-4), IL-13 and platelet-derived growth factor  
 CC (PDGF). Stat6b when compared Stat6 is an attenuated regulator of gene  
 CC transcription. Stat6c is a dominant negative regulator of gene  
 CC transcription.

SQ Sequence 18 AA;

Query Match 27.5%; Score 33; DB 20; Length 18;  
 Best Local Similarity 43.8%; Pred. No. 1.7e+02;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 7 EDIYPS--NPTDDVS 20  
   |||:|     ||:|:  
Db 2 edifppllppteqdlt 17

## RESULT 11

AAG64195  
ID AAG64195 standard; Peptide; 20 AA.

XX AC AAG64195;

XX DT 22-OCT-2001 (first entry)

XX DE Protein sequence 2.

XX KW Mycoplasma genitalium; complementary peptide; ligand;  
XX KW protein-protein interaction; drug design.

XX OS Unidentified.

XX PN WO200142278-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB04778.

XX PR 13-DEC-1999; 99GB-0029466.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-514238/56.

XX PT Complementary peptide ligands as reagents and drugs for drug discovery  
XX PT programs and as lead ligands to facilitate drug design and development,  
XX PT are generated from microbial genome sequences

PS Disclosure; Page 12; 161pp; English.

XX The invention relates to a set of complementary peptide ligands  
CC generated from Mycoplasma genitalium genome sequences. These  
CC specific complementary peptides interact with their relevant target  
CC proteins encoded by the microbial genome. They are capable of  
CC antagonising or agonising specific interaction of a protein with  
CC another protein or receptor and are thus useful as reagents and drugs,  
CC and as lead ligands to facilitate drug design and development. They  
CC are useful as tools for functional genomic studies, reagents for the  
CC configuration of high-throughput screens, as a starting point for  
CC medicinal chemistry manipulation, for peptide mimetics and as  
CC therapeutic agents. The analysis and acquisition of peptide sequences  
CC facilitates understanding of protein-protein interactions. The method  
CC allows for analysis of an entire database at a time, thus overcoming  
CC sampling problems. The present sequence is provided as an example  
CC protein sequence in a description of the invention.

XX Sequence 20 AA;

Query Match 27.5%; Score 33; DB 22; Length 20;  
Best Local Similarity 47.1%; Pred. No. 1.9e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 YRTNPEDIYPSNPTDDDD 18

Db :||:| | | | |

3 frtsredstysgtdtfd 19

## RESULT 12

AAB99270

ID AAB99270 standard; Peptide; 20 AA.

XX AC AAB99270;

XX DT 11-SEP-2001 (first entry)

XX DE Peptide sequence 2.

XX KW Plant; peptide pesticide; peptide herbicide; agricultural research.

XX OS Unidentified.

XX PN WO200142279-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB04781.

XX PR 13-DEC-1999; 99GB-0029469.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-381629/40.

XX PT A set of peptide ligands for agricultural research and development of  
XX PT therapeutic agents comprise specific complementary peptides to proteins  
XX PT encoded by genes of plant genomes

XX PS Disclosure; Page 10; 201pp; English.

XX CC The present invention relates to a set of peptide ligands consisting of  
XX CC specific complementary peptides to proteins encoded by genes of plant  
XX CC genomes e.g. from Arabidopsis thaliana (see AAG83361-AAG84504). The  
XX CC peptides of the present invention are useful in an assay to identify a  
XX CC peptide, especially a peptide pesticide or herbicide. The peptides are  
XX CC also useful for tools for agricultural research and development. The  
XX CC present peptide was used to illustrate the present invention.

XX Sequence 20 AA;

Query Match 27.5%; Score 33; DB 22; Length 20;  
Best Local Similarity 47.1%; Pred. No. 1.9e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 YRTNPEDIYPSNPTDDDD 18

Db :||:| | | | |  
3 frtsredstysgtdtfd 19

## RESULT 13

AAW34298

ID AAW34298 standard; peptide; 11 AA.

XX AC AAW34298;

XX DT 14-MAY-1998 (first entry)

XX DE Kringle 5 peptide fragment.

XX KW Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;  
XX KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;  
XX KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;  
XX KW autoimmune disease; ocular disease; capillary proliferation; therapy;  
XX KW kringle 5 receptor.

XX OS Synthetic.

XX PH Key Location/Qualifiers

XX FT Modified-site 1 /note= "N-Ac-Tyr"

FT Modified-site 11 /note= "C-terminal amide"  
 XX WO9741824-A2.  
 PN 13-NOV-1997.  
 PD 05-MAY-1997; 97WO-US07700.  
 PF 03-APR-1997; 97US-0832087.  
 PR 03-MAY-1996; 96US-0643219.  
 XX (ABBO ) ABBOTT LAB.  
 PA Davidson DJ, Gubbins EJ, Wang J;  
 PI WPI; 1997-558670/51.  
 DR New kringle 5 peptide(s) and fusion proteins derived from  
 XX plasminogen - useful as anti-angiogenesis agents for treating  
 XX cancer, psoriasis, arthritis etc., including gene therapy  
 PS Example 5; Page 43; 78pp; English.  
 XX This sequence is synthetic a kringle 5 (K5) peptide homologous to human  
 CC plasminogen. K5 peptide fragments homologous to this sequence, are  
 CC anti-angiogenesis agents, specifically for treating or preventing cancer,  
 CC particularly primary or metastatic solid tumours, carcinomas, sarcomas,  
 CC lymphomas, haemangiomas. They can also be used for treating or preventing  
 CC psoriasis, arthritis, macular degeneration and diabetic retinopathy. The  
 CC fragments can also be used to treat autoimmune or ocular diseases,  
 CC capillary proliferation within atherosclerotic plaque; haemophilic  
 CC joints, wound granulation, ulcers etc., also as contraceptives that  
 CC inhibit ovulation and establishment of the placenta. K5 antiserum or  
 CC (antagonists) can be used similarly, optionally coupled to cytotoxic  
 CC agents. Antagonists may be used to induce angiogenesis, e.g. for wound  
 CC healing. The K5 peptides are also used to raise specific antibodies (Ab),  
 CC for diagnosis and for affinity purification of K5 receptors. The K5  
 CC receptors may then be expressed in tumour cells to increase their  
 CC response to the peptides or used for identification of smaller  
 CC antagonists. The Ab are used to detect/quantify the peptides in  
 CC biological samples. The K5 peptides (and K5 fusion proteins) selectively  
 CC inhibit proliferation of endothelial cells with low toxicity against  
 CC normal cells. Typically they have 800-times greater inhibitory activity  
 CC against bovine capillary cells in vitro than kringle 1-4 peptides.  
 XX Sequence 11 AA;  
 SQ  
 Query Match 26.7%; Score 32; DB 18; Length 11;  
 Best Local Similarity 55.6%; Pred. NO. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 YRTNPEDIY 10  
 Db 1 yttmprkly 9  
 RESULT 14  
 AAB01899  
 ID AAB01899 standard; peptide; 11 AA.  
 XX AAB01899;  
 AC  
 XX 18-SEP-2000 (first entry)  
 DT Human plasminogen kringle 5 peptide fragment #5.  
 XX Plasminogen; human; kringle 5 domain; endothelial cell proliferation;  
 KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;  
 KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritic;  
 KW antiangiogenic; cancer; tumour; autoimmune disease.  
 XX

OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl moiety"  
 FT Modified-site 11  
 FT /note= "C-terminal amide moiety"  
 XX  
 PN US6057122-A.  
 XX 02-MAY-2000.  
 PD 05-MAY-1997; 97US-0851350.  
 PF 03-MAY-1996; 96US-0643219.  
 PR 03-APR-1997; 97US-0832087.  
 XX (ABBO ) ABBOTT LAB.  
 PA Davidson DJ;  
 PI WPI; 2000-349573/30.  
 DR Preparation of Kringle five peptide fragment for treating various  
 XX disorders such as angiogenic, Ocular, skin diseases and cancer,  
 PT involves mixing mammalian plasminogen and elastase followed by  
 PT incubation and isolation -  
 XX Example 5; Column 36; 48pp; English.  
 XX The invention relates to a method of preparing plasminogen kringle 5  
 CC peptide fragments. The method comprises mixing mammalian plasminogen and  
 CC elastase in the ratio 1:100-1:300, followed by incubating and isolating  
 CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and  
 CC endothelial cell proliferation and migration. The peptides are useful  
 CC for treating angiogenic diseases, primary and metastatic solid tumours  
 CC and carcinomas of various organs such as breast, genital tract,  
 CC endocrine glands, skin, tumours of the brain and eyes and solid tumours  
 CC arising from haematopoietic malignancies such as leukaemias and  
 CC lymphomas. They are also used for the prophylaxis of various autoimmune  
 CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases  
 CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber  
 CC Syndrome), diseases caused by excessive or abnormal stimulation of  
 CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases  
 CC which have angiogenesis as a pathologic consequence (e.g., cat scratch  
 CC disease and ulcers). The peptides are also useful as a birth control  
 CC agent which inhibits ovulation and establishment of the placenta.  
 CC Sequences AAB01888, AAB01889 and AAB01895-801905 represent human  
 CC plasminogen kringle 5-derived peptides synthesised and used in  
 CC exemplifications of the invention.  
 XX Sequence 11 AA;  
 SQ  
 Query Match 26.7%; Score 32; DB 21; Length 11;  
 Best Local Similarity 55.6%; Pred. NO. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 YRTNPEDIY 10  
 Db 1 yttmprkly 9  
 RESULT 15  
 AAB92094  
 ID AAB92094 standard; peptide; 11 AA.  
 XX AAB92094;  
 AC  
 XX 22-JUN-2001 (first entry)  
 DT Laminin fragment SEQ ID NO:1270.  
 XX

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidyl; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO200069900-A2.  
 XX 23-NOV-2000.  
 XX 17-MAY-2000; 2000WO-US13576.  
 XX 17-MAY-1999; 99US-0134406.  
 PR 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0159783.  
 XX (CONJ-) CONJUCHEM INC.  
 PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 PI WPI; 2001-112059/12.  
 XX  
 XX Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity  
 PT  
 XX  
 PS Disclosure; Page 611; 733pp; English.  
 XX  
 CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 11 AA;

Query Match 26.7%; Score 32; DB 22; Length 11;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRTNPEDIY 10  
 Db 1 ytnprkly 9

Search completed: February 21, 2002, 16:42:57  
 Job time: 231 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:40:46 ; Search time 31.57 Seconds  
(without alignments)  
14.969 Million cell updates/sec

Title: US-08-753-851-5  
Perfect score: 120  
Sequence: 1 CYRTNPEDIYPSNPTDDVSS 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 114342

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	21	2 US-08-143-311B-5	Sequence 5, Appli
2	35	29.2	12	4 US-08-938-595-5	Sequence 5, Appli
3	35	29.2	12	4 US-08-727-153-5	Sequence 5, Appli
4	34	28.3	18	1 US-07-729-099-15	Sequence 15, Appl
5	34	28.3	18	1 US-08-257-392-15	Sequence 15, Appl
6	34	28.3	18	3 US-08-770-035-15	Sequence 15, Appl
7	32	26.7	16	1 US-08-116-733-4	Sequence 4, Appli
8	32	26.7	16	1 US-08-469-615-3	Sequence 3, Appli
9	32	26.7	16	1 US-08-466-763-3	Sequence 3, Appli
10	32	26.7	16	2 US-08-411-142A-3	Sequence 3, Appli
11	31	25.8	9	2 US-08-318-856A-53	Sequence 53, Appl
12	31	25.8	15	2 US-08-857-068-3	Sequence 3, Appli
13	31	25.8	19	1 US-08-520-977A-8	Sequence 8, Appli
14	30	25.0	8	4 US-08-938-595-4	Sequence 4, Appli
15	30	25.0	8	4 US-08-727-153-4	Sequence 4, Appli
16	30	25.0	15	3 US-08-630-916A-105	Sequence 105, App
17	30	25.0	15	4 US-08-602-999A-311	Sequence 311, App
18	30	25.0	16	6 5378805-4	Patent No. 5378805
19	30	25.0	19	1 US-07-596-081A-28	Sequence 28, Appl
20	30	25.0	19	5 PCT-US96-10905-18	Sequence 18, Appl
21	30	25.0	20	1 US-08-732-751-10	Sequence 10, Appl
22	30	25.0	20	4 US-08-612-973-86	Sequence 86, Appl
23	30	25.0	20	4 US-08-927-597-86	Sequence 86, Appl
24	29.5	24.6	14	4 US-08-557-006B-19	Sequence 19, Appl
25	29	24.2	12	2 US-08-934-222-25	Sequence 25, Appl
26	29	24.2	12	2 US-08-933-402-25	Sequence 25, Appl
27	29	24.2	12	2 US-09-207-621-25	Sequence 25, Appl

28 29 24.2 12 2 US-08-532-818-25 Sequence 25, Appl  
29 24.2 12 3 US-09-231-797-25 Sequence 25, Appl  
30 29 24.2 12 3 US-08-934-224-25 Sequence 25, Appl  
31 29 24.2 12 3 US-08-933-843-25 Sequence 25, Appl  
32 29 24.2 12 4 US-08-934-223-25 Sequence 25, Appl  
33 29 24.2 12 4 US-09-413-492-25 Sequence 25, Appl  
34 29 24.2 16 1 US-07-663-413-21 Sequence 21, Appl  
35 29 24.2 16 1 US-08-055-530-21 Sequence 21, Appl  
36 29 24.2 18 2 US-08-616-844-50 Sequence 50, Appl  
37 29 24.2 18 2 US-08-599-654-50 Sequence 50, Appl  
38 29 24.2 18 2 US-08-759-599-5 Sequence 5, Appl  
39 29 24.2 18 2 US-09-017-205-31 Sequence 31, Appl  
40 29 24.2 18 3 US-08-944-868A-50 Sequence 50, Appl  
41 29 24.2 18 3 US-08-944-423A-50 Sequence 50, Appl  
42 29 24.2 18 3 US-08-944-496-50 Sequence 50, Appl  
43 29 24.2 20 1 US-08-466-033-85 Sequence 85, Appl  
44 29 24.2 20 1 US-08-466-033-89 Sequence 89, Appl  
45 29 24.2 20 2 US-08-444-733-85 Sequence 85, Appl

## ALIGNMENTS

RESULT 1  
US-08-143-311B-5  
; Sequence 5, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHIVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-5

Qy	11	PSNP	DDD	18
		1:	1111	
Db	3	PAFP	DDD	10

Qy	11	PSNP	DDD	18
		1:	1111	
Db	3	PAFP	DDD	10

```

;
; APPLICANT: pasamontes, Luis
;
; TITLE OF INVENTION: Coccidiosis vaccine
;
; NUMBER OF SEQUENCES: 15
;
; CORRESPONDENCE ADDRESS:

```



ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/729,099  
FILING DATE: 19910712  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Catherine R  
REGISTRATION NUMBER: 34,240  
REFERENCE/DOCKET NUMBER: 8514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
\* INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FRAGMENT TYPE: C-terminal  
ORIGINAL SOURCE:  
ORGANISM: Eimeria tenella  
US-07-729-099-15

Query Match 28.3%; Score 34; DB 1; Length 18;  
Best Local Similarity 40.0%; Pred. No. 45;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 TNPEDIYPSNPTDDD 18  
| | : | | | |  
Db 1 TQPDTEFRSGPGDDE 15

RESULT 5  
US-08-257-392-15  
Sequence 15, Application US/08257392  
Patent No. 5688513  
GENERAL INFORMATION:  
APPLICANT: Binger, Mary-Helen  
APPLICANT: Pasamontes, Luis  
TITLE OF INVENTION: Coccidiosis Vaccines  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/257,392  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/729,099  
FILING DATE: 12-JUL-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Roseman, Catherine R  
REGISTRATION NUMBER: 34,240  
REFERENCE/DOCKET NUMBER: 8514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FRAGMENT TYPE: C-terminal  
ORIGINAL SOURCE:  
ORGANISM: Eimeria tenella  
US-08-257-392-15

Query Match 28.3%; Score 34; DB 1; Length 18;  
Best Local Similarity 40.0%; Pred. No. 45;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 TNPEDIYPSNPTDDD 18  
| | : | | | |  
Db 1 TQPDTEFRSGPGDDE 15

RESULT 6  
US-08-770-035-15  
Sequence 15, Application US/08770035  
Patent No. 6008342  
GENERAL INFORMATION:  
APPLICANT: Binger, Mary-Helen  
APPLICANT: Pasamontes, Luis  
TITLE OF INVENTION: Coccidiosis Vaccines  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/729,099  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Catherine R  
REGISTRATION NUMBER: 34,240  
REFERENCE/DOCKET NUMBER: 8514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FRAGMENT TYPE: C-terminal  
ORIGINAL SOURCE:  
ORGANISM: Eimeria tenella

US-08-770-035-15

Query Match 28.3%; Score 34; DB 3; Length 18;  
Best Local Similarity 40.0%; Pred. No. 45;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 TNPEDYPSNPTDD 18  
DB 1 TOPDTEFRSGDDE 15

RESULT 7

US-08-116-733-4

; Sequence 4, Application US/08116733

; Patent No. 5516632

; GENERAL INFORMATION:

; APPLICANT: PALKER, Thomas J.

; APPLICANT: HAYNES, Barton F.

; TITLE OF INVENTION: SYNTHETIC PEPTIDES

; NUMBER OF SEQUENCES: 46

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON &amp; VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/116,733

; FILING DATE: 07-SEP-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: WILSON, MARY J.

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 1579-33

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; TELEX: 200797 NIXN UR

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 16 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-116-733-4

Query Match 26.7%; Score 32; DB 1; Length 16;  
Best Local Similarity 35.7%; Pred. No. 77;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 CYRTNPEDIYPSNP 14  
DB 1 CLNTEPSQLPPTAP 14

RESULT 8

US-08-469-615-3

; Sequence 3, Application US/08469615

; Patent No. 5622703

; GENERAL INFORMATION:

; APPLICANT: Berzofsky, Jay A.

; APPLICANT: Kurata, Akihiko

; TITLE OF INVENTION: Immunodominant Sites of HTLV-I Envelope

; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 19

;

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch &amp; Birch, LLP

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,615

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/243,118

FILING DATE: 16-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Svensson, Leonard R.

REGISTRATION NUMBER: 30330

REFERENCE/DOCKET NUMBER: 1173-500P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-205-8000

TELEFAX: 703-205-8050

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

ORGANISM: htlv-I

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..16

OTHER INFORMATION: /label= peptide\_sp3

US-08-469-615-3

Query Match 26.7%; Score 32; DB 1; Length 16;  
Best Local Similarity 35.7%; Pred. No. 77;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 CYRTNPEDIYPSNP 14  
DB 1 CLNTEPSQLPPTAP 14

RESULT 9

US-08-466-763-3

; Sequence 3, Application US/08466763

; Patent No. 5695762

; GENERAL INFORMATION:

; APPLICANT: Berzofsky, Jay A.

; APPLICANT: Kurata, Akihiko

; TITLE OF INVENTION: Immunodominant Sites of HTLV-I Envelope

; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch &amp; Birch, LLP

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,763  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,118  
FILING DATE: 16-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30330  
REFERENCE/DOCKET NUMBER: 1173-497P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: htlv-I  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..16  
OTHER INFORMATION: /label= peptide\_sp3  
US-08-466-763-3

Query Match 26.7%; Score 32; DB 1; Length 16;  
Best Local Similarity 35.7%; Pred. No. 77;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CYTNPEDIYPSNP 14  
| | | | | | | | | | | | | | | |  
Db 1 CLNTEPSQLPPTAP 14

RESULT 10  
US-08-411-142A-3  
Sequence 3, Application US/08411142A  
Patent No. 5882853  
GENERAL INFORMATION:  
APPLICANT: Berzofsky, Jay A.  
APPLICANT: Kurata, Akihiko  
TITLE OF INVENTION: METHOD OF EVALUATING HTLV-I-SPECIFIC  
TITLE OF INVENTION: T CELL RESPONSES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/411.142A  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/243,118  
FILING DATE: 16-MAY-1994  
APPLICATION NUMBER: 07/401,411  
FILING DATE: 01-SEP-1989

ATTORNEY/AGENT INFORMATION:  
NAME: Wetherelli, Jr., Ph.D, John R.  
REGISTRATION NUMBER: 31,678  
REFERENCE/DOCKET NUMBER: 08830/026002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
IMMEDIATE SOURCE:  
CLONE: HTLV-I  
US-08-411-142A-3

Query Match 26.7%; Score 32; DB 2; Length 16;  
Best Local Similarity 35.7%; Pred. No. 77;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CYTNPEDIYPSNP 14  
| | | | | | | | | | | | | | | |  
Db 1 CLNTEPSQLPPTAP 14

RESULT 11  
US-08-318-856A-53  
Sequence 53, Application US/08318856A  
Patent No. 5972351  
GENERAL INFORMATION:  
APPLICANT: Adrian V.S. Hill, et al.  
TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-  
TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
TITLE OF INVENTION: ANTIGENS (AS AMENDED)  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1+  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,856A  
FILING DATE: October 3, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 92 08 068.8  
FILING DATE: April 3, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 92 17 704.7  
FILING DATE: August 20, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB93/00711  
FILING DATE: April 5, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng  
REGISTRATION NUMBER: 40,949  
REFERENCE/DOCKET NUMBER: 263-PPTR157705  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 721-8200  
TELEFAX: (202) 721-8250  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acid residues  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-318-856A-53

Query Match 25.8%; Score 31; DB 2; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.6e-05;  
Matches 5; Conservative 0; Mismatches 2; Indels 0;

QY 11 PSNPTDD 17  
| | | | |  
Db 2 PKNPEDD 8

RESULT 12  
US-08-857-068-3  
; Sequence 3, Application US/08857068  
; Patent No. 5866682  
; GENERAL INFORMATION:  
; APPLICANT: McIntosh, J. Michael  
; APPLICANT: Cartier, G. Edward  
; APPLICANT: Yoshikami, Doju  
; APPLICANT: Luo, Sign  
; APPLICANT: Olivera, Baldomero M.  
; TITLE OF INVENTION: CONOPEPTIDES AUIA, AU1B AND AU1C  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
; STREET: 1201 New York Avenue, Suite 1000  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: US  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/857,068  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ihnen, Jeffrey L.  
; REGISTRATION NUMBER: 28,957  
; REFERENCE/DOCKET NUMBER: 24260-121388  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-962-4810  
; TELEFAX: 202-962-8300  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Conus aulicus  
; FEATURE:  
; NAME/KEY: Disulfide-bond  
; LOCATION: 2..8  
; FEATURE:  
; NAME/KEY: Disulfide-bond  
; LOCATION: 3..15  
US-08-857-068-3

Query Match 25.8%; Score 31; DB 2; Length 15;  
Best Local Similarity 57.1%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0;

QY 1 CYRTNPE 7  
| | | | |  
Db 8 CFATNPD 14

RESULT 13  
US-08-520-977A-8  
; Sequence 8, Application US/08520977A  
; Patent No. 5659018  
; GENERAL INFORMATION:  
; APPLICANT: Berndt, Michael C.  
; APPLICANT: Dunlop, Lindsay  
; APPLICANT: Andrews, Robert  
; APPLICANT: DeLuca, Mariagrazia  
; TITLE OF INVENTION: MOCARHAGIN, A NOVEL COBRA VENOM  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/520,977A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; REFERENCE/DOCKET NUMBER: GI5239  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-520-977A-8

Query Match 25.8%; Score 31; DB 1; Length 19;  
Best Local Similarity 40.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 TNPDIYPSNPTDDD 18  
| | | | |  
Db 5 TDLVDYYPEDTEGD 19

RESULT 14  
US-08-938-595-4  
; Sequence 4, Application US/08938595  
; Patent No. 6197946  
; GENERAL INFORMATION:  
; APPLICANT: Cottingham, Ian R.  
; APPLICANT: Garner, Ian  
; TITLE OF INVENTION: Peptide Production as Fusion Protein in  
; TITLE OF INVENTION: Transgenic Mammal Milk  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600

CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,595  
FILING DATE: (Herewith)  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/727,153  
FILING DATE: 08-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/718,523  
FILING DATE: 08-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB95/00769  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9406974.7  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fleshner, Raz E.  
REGISTRATION NUMBER: 34,331  
REFERENCE/DOCKET NUMBER: 0623.0460003  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-938-595-4

Query Match 25.0%; Score 30; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 PTDDD 18  
Db 2 PTDDD 6

RESULT 15  
US-08-727-153-4  
Sequence 4, Application US/08727153  
Patent No. 6211427  
GENERAL INFORMATION:  
APPLICANT: Cottingham, Ian R.  
TITLE OF INVENTION: Peptide Production as Fusion Protein in  
TITLE OF INVENTION: Transgenic Mammal Milk  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/727,153  
FILING DATE: 08-OCT-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/718,523  
FILING DATE: 08-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB95/00769  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9406974.7  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fleshner, Raz E.  
REGISTRATION NUMBER: 34,331  
REFERENCE/DOCKET NUMBER: 0623.0460001  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-727-153-4

Query Match 25.0%; Score 30; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 PTDDD 18  
Db 2 PTDDD 6

Search completed: February 21, 2002, 16:43:35  
Job time: 169 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:44:15 ; Search time 33.6 Seconds  
(without alignments)  
47.609 Million cell updates/sec

Title: US-08-753-851-6

Perfect score: 123

Sequence: 1 CTVHPIDEDSPWITDTPRI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4046

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	30	24.4	18	2	A35678
2	28	22.8	21	2	S32883
3	27	22.0	16	2	S66613
4	27	22.0	16	2	A20190
5	27	22.0	18	1	A58589
6	27	22.0	19	2	S60633
7	26	21.1	16	2	B23692
8	26	21.1	20	2	S10876
9	25	20.3	11	2	JQ2307
10	25	20.3	15	2	S29207
11	25	20.3	18	2	A39040
12	25	20.3	21	2	S08590
13	24.5	19.9	20	2	A44773
14	24	19.5	8	2	A61328
15	24	19.5	11	2	PS0259
16	24	19.5	13	2	A60856
17	24	19.5	15	2	FX0031
18	24	19.5	17	2	A44799
19	24	19.5	19	2	B26930
20	24	19.5	19	2	S63489
21	24	19.5	19	2	B46592
22	24	19.5	20	2	S77981
23	24	19.5	20	2	S63490
24	23.5	19.1	17	2	A42920
25	23	18.7	9	2	JS0302
26	23	18.7	9	2	A60320
27	23	18.7	15	2	A56863
28	23	18.7	15	2	PA0088
29	23	18.7	19	2	PC2326

cytotoxic T-lympho  
T-cell receptor J-  
pev-kinin 1 - pena  
leucopyrokinin - M  
substance P-like p  
kassinin-like pept  
kassinin-like pept  
amine oxidase (cop  
acidic proline-ric  
Ig heavy chain CDR  
neural cell adhesi  
hypothetical prote  
collecting duct wa  
Ig heavy chain CDR  
protein kinase (EC  
agrin - electric r

#### ALIGNMENTS

RESULT 1

A35678

hypothetical protein (proenkephalin 5' region) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 05-Nov-1999

C;Accession: A35678

R;Kilpatrick, D.L.; Zinn, S.A.; Fitzgerald, M.; Higuchi, H.; Sabol, S.L.; Meyerhardt, Mol. Cell. Biol. 10, 3717-3726, 1990

A;Title: Transcription of the rat and mouse proenkephalin genes is initiated at distal

A;Reference number: A35678; MUID:90287163

A;Accession: A35678

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-18 <KIL>

A;Cross-references: GB:M55181; NID:g201032; PIDN:AAA40127.1; PID:g201033

Query Match 24.4%; Score 30; DB 2; Length 18;  
Best Local Similarity 80.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 9 EDSWP 13

Db 6 QDSPW 10

RESULT 2

S32883

DNA gyrase chain A - Streptomyces spheroides (fragment)

C;Species: Streptomyces spheroides

C;Date: 08-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 22-Oct-1999

C;Accession: S32883; S29684

R;Thiara, A.S.; Cundliffe, E.

Mol. Microbiol. 8, 495-506, 1993

A;Title: Expression and analysis of two gyrB genes from the novobiocin producer, Stre

A;Reference number: S32881; MUID:93316846

A;Accession: S32883

A;Molecule type: DNA

A;Residues: 1-21 <THI>

A;Cross-references: EMBL:Z17305; NID:g47535; PIDN:CAA78953.1; PID:g47537

Query Match 22.8%; Score 28; DB 2; Length 21;  
Best Local Similarity 28.6%; Pred. No. 6.2e+02;  
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 8 DEDSPWITDTPRI 21

Db 3 DENTPMPEEYPAV 16

RESULT 3

S66613  
 protein p12E - Friend murine leukemia virus (fragments)  
 C:Species: Friend murine leukemia virus  
 C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997  
 C:Accession: S66613  
 R:Hensel, J.; Hintz, M.; Karas, M.; Linder, D.; Stahl, B.; Geyer, R.  
 Eur. J. Biochem. 232, 373-380, 1995  
 A:Title: Localization of the palmitoylation site in the transmembrane protein p12E of Friend murine leukemia virus  
 A:Reference number: S66613; MUID:96035869  
 A:Accession: S66613  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9;10-16 <REN>

Query Match 22.0%; Score 27; DB 2; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 SPWIT 15  
 |||  
 Db 10 SPWFT 14

RESULT 4  
 A20190  
 hypodermin B - early cattle grub (fragment)  
 C:Species: Hypoderma lineatum (early cattle grub)  
 C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 18-Jun-1993  
 C:Accession: A20190  
 R:Lecroisey, A.; Tong, N.T.; Keil, B.  
 Eur. J. Biochem. 134, 261-267, 1983  
 A:Title: Hypodermin B, a trypsin-related enzyme from the insect Hypoderma lineatum.  
 A:Reference number: A20190; MUID:83261874  
 A:Accession: A20190  
 A:Molecule type: protein  
 A:Residues: 1-16 <LEC>

Query Match 22.0%; Score 27; DB 2; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 EDSPPW 13  
 ||||  
 Db 10 EDPFW 14

RESULT 5  
 A58589  
 alpha-conotoxin EI - cone shell (Conus ermineus)  
 C:Species: Conus ermineus (ermine cone)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A58589  
 R:Martinez, J.S.; Olivera, B.M.; Gray, W.R.; Craig, A.G.; Groebe, D.R.; Abramson, S.N.;  
 Biochemistry 34, 14519-14526, 1995  
 A:Title: alpha-Conotoxin EI, a new nicotinic acetylcholine receptor antagonist with novel pharmacological properties  
 A:Reference number: A58589; MUID:96062516  
 A:Accession: A58589  
 A:Molecule type: protein  
 A:Residues: 1-18 <MAR>  
 A>Note: sequence confirmed by chemical synthesis  
 C:Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynaptic antagonist of the alpha-conotoxin  
 C:Superfamily: alpha-conotoxin  
 C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline; post-synaptic  
 F:3/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:4-10, 5-18/Disulfide bonds: #status experimental  
 F:18/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 22.0%; Score 27; DB 1; Length 18;  
 Best Local Similarity 35.7%; Pred. No. 7.2e+02;  
 Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTVHPIDEDSPWI 14  
 |||  
 Db 4 CCYHPTCNSNPQI 17

RESULT 6  
 S60633  
 H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - brine shrimp mitochondrion (strain 8)  
 C:Species: mitochondrion Artemia sp. (brine shrimp)  
 A:Variety: strain La Mata  
 C:Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-Dec-1999  
 C:Accession: S60633  
 R:Perez, M.L.; Valverde, J.R.; Batuecas, B.; Amat, F.; Marco, R.; Garesse, R.  
 J. Mol. Evol. 38, 156-168, 1994  
 A:Title: Speciation in the Artemia genus: mitochondrial DNA analysis of bisexual and asexual populations  
 A:Reference number: S60633; MUID:94223692  
 A:Accession: S60633  
 A:Molecule type: DNA  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A:Residues: 1-19 <PER>

A:Cross-references: EMBL:X67263; NID:g11210; PIDN:CAA47685.1; PID:g11211  
 A:Experimental source: strain La Mata  
 A>Note: the source is designated as Artemia parthenogenetica  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992  
 C:Genetics:  
 A:Gene: ATP8  
 A:Genome: mitochondrion  
 A:Genetic code: SEC4  
 C:Superfamily: H+-transporting ATP synthase protein 8  
 C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion;

Query Match 22.0%; Score 27; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PWIT 15  
 ||||  
 Db 8 PWIT 11

RESULT 7  
 B23692  
 transcription factor chain All, CCAAT-binding - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 30-Sep-1993  
 C:Accession: B23692  
 R:Vuorio, T.; Maity, S.N.; de Crombrughe, B.  
 J. Biol. Chem. 265, 22480-22486, 1990  
 A:Title: Purification and molecular cloning of the "A" chain of a rat heteromeric CCAAT-binding factor  
 A:Reference number: A23692; MUID:91093096  
 A:Accession: B23692  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-16 <VUU>  
 A:Cross-references: GB:J05701

Query Match 21.1%; Score 26; DB 2; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 8.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 EDSPPW 13  
 ||||  
 Db 9 EPPSW 13

RESULT 8  
 S10876  
 hypothetical protein - human  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 18-Nov-1994



C:Accession: S10876  
 R:Assouline, Z.; Kerbirou-Nabias, D.M.; Pietu, G.; Thomas, N.; Bahnak, B.R.; Meyer, D.  
 Biochem. Biophys. Res. Commun. 153, 1159-1166, 1988  
 A:Title: The human gene for von Willebrand factor. Identification of repetitive Alu sequences  
 A:Reference number: S10876; MUID:88268889  
 A:Accession: S10876  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-20 <ASS>  
 A:Cross-references: EMBL:X07258

Query Match 21.1%; Score 26; DB 2; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 SPWIT 15  
 Db 16 SPWAT 20

RESULT 9  
 JQ2307  
 hypotetical 1.5K protein - tomato chloroplast (strain Toko)  
 C:Species: chloroplast Lycopersicon esculentum (tomato)  
 C:Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 23-Mar-1995  
 C:Accession: JQ2307  
 R:Kawagoe, Y.; Kikuta, Y.  
 Theor. Appl. Genet. 81, 13-20, 1991  
 A:Title: Chloroplast DNA evolution in potato (Solanum tuberosum L.).  
 A:Reference number: JQ2306  
 A:Accession: JQ2307  
 A:Molecule type: DNA  
 A:Residues: 1-11 <KAW>  
 A:Experimental source: strain Toko  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast

Query Match 20.3%; Score 25; DB 2; Length 11;  
 Best Local Similarity 57.1%; Pred. No. 8e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 DEBSPWI 14  
 Db 5 DEFTWI 11

RESULT 10  
 S29207  
 avenin gamma-4 - oat (fragment)  
 N:Alternate names: CIP-1; coeliac immunoreactive protein 1  
 C:Species: Avena sativa (oat)  
 C:Date: 19-Mar-1997 #sequence\_revision 24-Jul-1998 #text\_change 24-Jul-1998  
 C:Accession: S29207  
 R:Rocher, A.; Collilla, F.; Ortiz, M.L.; Mendez, E.  
 FEBS Lett. 310, 37-40, 1992  
 A:Title: Identification of the three major coeliac immunoreactive proteins and one alpha  
 A:Reference number: S29207; MUID:92405739  
 A:Accession: S29207  
 A:Molecule type: protein  
 A:Residues: 1-15 <ROC>  
 A:Experimental source: endosperm  
 C:Superfamily: gliadin  
 C:Keywords: prolamin; seed

Query Match 20.3%; Score 25; DB 2; Length 15;  
 Best Local Similarity 33.3%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PIPDESPW 13

Db 7 PYEQOQPF 15

RESULT 11  
 A39040  
 caldesquestrin, cardiac muscle - dog (fragments)  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 23-Feb-1997  
 C:Accession: A39040  
 R:Calal, S.E.; Jones, L.R.  
 J. Biol. Chem. 266, 391-398, 1991  
 A:Title: Phosphorylation of cardiac and skeletal muscle caldesquestrin isoforms by cas  
 A:Reference number: A39040; MUID:91093153  
 A:Accession: A39040  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-18 <CAL>  
 C:Keywords: cardiac muscle; heart; phosphoprotein

Query Match 20.3%; Score 25; DB 2; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 IPDED 10  
 Db 2 IPDDD 6

RESULT 12  
 S08590  
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Debney's tobacco chloroplast (C:Species: chloroplast Nicotiana debneyi (Debney's tobacco)  
 C:Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 11-Jun-1999  
 C:Accession: S08590  
 R:Hayashida, N.; Matsubayashi, T.; Shinozaki, K.; Sugiyura, M.; Inoue, K.; Hiyama, T.  
 Curr. Genet. 12, 247-250, 1987  
 A:Title: The gene for the 9 kd polypeptide, a possible apoprotein for the iron-sulfur  
 A:Reference number: S07170; MUID:88210537  
 A:Accession: S08590  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-21 <HAY>  
 A:Cross-references: EMBL:X05881; NID:g11790; PIDN:CAA29305.1; PID:g11793  
 C:Genetics:  
 A:Gene: ndhd  
 A:Genome: chloroplast  
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
 C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 20.3%; Score 25; DB 2; Length 21;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PWIT 15  
 Db 14 PWLT 17

RESULT 13  
 A44773  
 pollen allergen I - Japanese cedar (fragment)  
 C:Species: Cryptomeria japonica (Japanese cedar)  
 C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 11-Jan-2000  
 C:Accession: A44773  
 R:Taniai, M.; Ando, S.; Usui, M.; Kurimoto, M.; Sakaguchi, M.; Inouye, S.; Matuhasi,  
 FEBS Lett. 239, 329-332, 1988  
 A:Title: N-terminal amino acid sequence of a major allergen of Japanese cedar pollen  
 A:Reference number: A44773; MUID:89031257  
 A:Accession: A44773  
 A>Status: preliminary

A:Molecule type: protein  
A:Residues: 1-20 <FAN>  
C:Superfamily: pectate lyase LAT59  
C:Keywords: pollen

Query Match 19.9%; Score 24.5; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. NO. 1.9e+03;  
Matches 7; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 4 HPIPEDSPWITDS 17  
:|| || ||  
Db 2 NPI---DSXWRGDS 12

## RESULT 14

A61328  
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)  
C:Species: Balaenoptera acutorostrata (minke whale, lesser rorqual)  
C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 07-May-1999  
C:Accession: A61328  
R:Brictoux-Gregoire, S.; Schyns, R.; Florkin, M.; Emmens, M.; Welling, G.W.; Beintema, J.  
Biochim. Biophys. Acta 386, 244-255, 1975  
A:Title: N-terminal amino acid sequence of trypsinogen from the lesser rorqual, Balaenoptera acutorostrata (minke whale) cDNA.  
A:Reference number: A61328; MUID:75146765  
A:Accession: A61328  
A:Molecule type: protein  
A:Residues: 1-8 <BRI>  
C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen  
F:1-8/Domain: activation peptide #status experimental <APT>

Query Match 19.5%; Score 24; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. NO. 2.2e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PIPDED 10  
|| ||  
Db 2 PIDDDD 7

## RESULT 15

PS0259  
39K protein 3225 - rice (strain Nihonbare) (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995  
C:Accession: PS0259  
R:Tsuigita, A.; Kamo, M.  
submitted to JIPID, April 1993  
A:Reference number: PS0209  
A:Accession: PS0259  
A:Molecule type: protein  
A:Residues: 1-11 <TSU>  
A:Experimental source: callus  
C:Comment: molecular weight 39K, pI 5.7.

Query Match 19.5%; Score 24; DB 2; Length 11;  
Best Local Similarity 50.0%; Pred. NO. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 EDSPPWITD 16  
|| || ||  
Db 4 EDGPVAD 11

Search completed: February 21, 2002, 16:44:16  
Job time: 190 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:52:20 ; Search time 20.15 Seconds  
(without alignments)  
38.211 Million cell updates/sec

Title: US-08-753-851-6  
Perfect score: 123  
Sequence: 1 CTVHPIDEDSPWITDTPRI 21

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1160

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	24.4	11	1	TINL_HOPTI
2	28	22.8	21	1	GYRA_STRSH
3	27	22.0	18	1	CXAL_CONER
4	24	19.5	16	1	TRYP_FELCA
5	24	19.5	20	1	COXA_THUOB
6	23	18.7	8	1	LMT2_LOCMI
7	23	18.7	16	1	PPAC_BACME
8	22	17.9	8	1	LPK_LEUMA
9	22	17.9	11	1	PKC1_CARMO
10	22	17.9	11	1	TKN1_PSEGU
11	22	17.9	11	1	TKN2_PSEGU
12	22	17.9	11	1	TKN4_PSEGU
13	22	17.9	13	1	SODM_CANFA
14	22	17.9	19	1	FIBB_PIG
15	22	17.9	21	1	ECD1_LYMDI
16	22	17.9	21	1	TRYP_PROAT
17	21	17.1	13	1	ECDE_LYMDI
18	21	17.1	13	1	LPAA_FORGI
19	21	17.1	15	1	UEI5_HORVU
20	21	17.1	17	1	LCK_RAT
21	21	17.1	19	1	DHAB_COMTE
22	21	17.1	20	1	CATL_FASHE
23	21	17.1	20	1	COXB_THUOB
24	21	17.1	20	1	LECB_TRIHO
25	21	17.1	20	1	RECX_AZOVI
26	21	17.1	20	1	SCIB_BPT5
27	20	16.3	11	1	TIN4_HOPTI
28	20	16.3	12	1	TIN2_HOPTI
29	20	16.3	12	1	TIN3_HOPTI
30	20	16.3	12	1	ULAL_MOUSE
31	20	16.3	15	1	IRBP_CRISP
32	20	16.3	17	1	GAST_WACMU
33	20	16.3	17	1	PROP_SALTY

34	20	16.3	19	1	LPRM_STAAR	P03063 staphylococ
35	20	16.3	19	1	TRPB_KLEAE	P14552 klebsiella
36	20	16.3	21	1	SRTD_ATREN	P13211 atractaspis
37	19.5	15.9	16	1	FOR2_MYRGU	P81437 myrmecia gu
38	19	15.4	9	1	MGMT_BOVIN	P29177 bos taurus
39	19	15.4	10	1	ANG1_BOTJA	Q10581 bothrops ja
40	19	15.4	10	1	ANGT_BOVIN	P01017 bos taurus
41	19	15.4	10	1	ANGT_CHICK	P01018 gallus gall
42	19	15.4	10	1	LPK2_LOCMI	P41488 locusta mig
43	19	15.4	10	1	TOPI_SALTY	P40686 salmoneilla
44	19	15.4	11	1	ANGT_CRIGE	P09037 crinia geor
45	19	15.4	12	1	TKN1_KASMA	P08613 kassina mac

ALIGNMENTS

RESULT 1					
TINL_HOPTI					
ID	TINL_HOPTI	STANDARD;	PRT;	11 AA.	
AC	P82651;				
DT	20-AUG-2001 (Rel. 40, Created)				
DT	20-AUG-2001 (Rel. 40, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	TIGERININ-1.				
OS	Hoplobatrachus tigerinus (Rana tigerina).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae;				
OC	Hoplobatrachus.				
OX	NCBI_TaxID=103373;				
RN	[1]				
RP	SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.				
RC	TISSUE=SKIN;				
RX	PubMed=11031261;				
RA	Purna Sai K., Jagannadham M.V., Vairamani M., Raju N.P.,				
RA	Devi A.S., Nagaraj R., Sitaram N.;				
RT	"Tigerinins: novel antimicrobial peptides from the Indian frog Rana				
RT	tigerina.";				
RL	J. Biol. Chem. 276:2701-2707(2001).				
CC	-I- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST B.SUBTILIS, E.COLI,				
CC	S.AUREUS, M.LUTEUS, P.PUTIDA AND S.CEREVISIAE.				
CC	-I- SUBCELLULAR LOCATION: SECRETED.				
CC	-I- TISSUE SPECIFICITY: SKIN.				
CC	-I- MASS SPECTROMETRY: MW=1342; METHOD=MALDI.				
KW	Amphibian skin; Antibiotic; Amidation.				
FT	DISULFID 2 10				
FT	MOD_RES 11 11				
FT	AMIDATION.				
SQ	SEQUENCE 11 AA; 1344 MW; A2087DC960476056 CRC64;				
Query Match 24.4%; Score 30; DB 1; Length 11;					
Best Local Similarity 71.4%; Pred. No. 82;					
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;					
Qy	1 CTVHPIP 7				
	:				
Db	2 CTMIPIP 8				
RESULT 2					
GYRA_STRSH					
ID	GYRA_STRSH	STANDARD;	PRT;	21 AA.	
AC	P50071;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	DNA GYRASE SUBUNIT A (EC 5.99.1.3) (FRAGMENT).				
GN	GYRA.				
OS	Streptomyces sp. sphaeroides.				
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
OC	Actinomycetales; Streptomycetaceae; Streptomyces.				
OX	NCBI_TaxID=29307;				
RN	[1]				

RP SEQUENCE FROM N.A.  
RC STRAIN-NCIB 11891;  
RX MEDLINE-93316846; PubMed-8392138;  
RA Thiara A.S., Cundliffe E.;  
RT "Expression and analysis of two gyrB genes from the novobiocin  
RT producer, Streptomyces sphaeroides";  
RL Mol. Microbiol. 8:495-506(1993).  
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE  
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED  
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING  
CC OF DOUBLE-STRANDED DNA.  
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA  
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE  
CC ENZYME FORMS AN A2B2 TETRAMER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z17305; CAA78953.1;  
DR Topoisomerase; Isomerase; DNA-binding.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2203 MW; 5E2F9DEA8DB83697 CRC64;

Query Match 22.8%; Score 28; DB 1; Length 21;  
Best Local Similarity 28.6%; Pred. No. 3.2e+02;  
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 8 DEDSPWTDTPRI 21  
||:|:|:|:|:  
Db 3 DENTPVMPVEVPV 16

## RESULT 3

ID CXAL\_CONER STANDARD; PRT; 18 AA.  
AC P50982;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE ALPHA-CONOTOXIN EI.  
OS Conus ermineus (Atlantic fish-hunting cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=55423;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE-96062516; PubMed-7578057;  
RA Martinez J.S., Olivera B.M., Gray W.R., Craig A.G., Groebe D.R.,  
RA Abranson S.N., McIntosh J.M.;  
RT "Alpha-conotoxin EI, a new nicotinic acetylcholine receptor  
RT antagonist with novel selectivity";  
RL Biochemistry 34:14519-14526(1995).  
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
CC INHIBIT THEM.  
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
KW Venom; Hydroxylation.  
FT DISULFID 4 10  
FT DISULFID 5 18  
FT MOD\_RES 3 3 HYDROXYLATION.  
FT MOD\_RES 18 18 AMIDATION.  
SQ SEQUENCE 18 AA; 2082 MW; 60A61A6C427A6B5E CRC64;

Query Match 22.0%; Score 27; DB 1; Length 18;  
Best Local Similarity 35.7%; Pred. No. 3.7e+02;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTVPHPDEDSPI 14  
|:|:|:|:  
Db 4 CCYHPTCNMNPQI 17

## RESULT 4

ID TRYP\_FELCA STANDARD; PRT; 16 AA.  
AC P81071;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE TRYPSIN PRECURSOR (EC 3.4.21.4) (FRAGMENT).  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Pancreas;  
RX MEDLINE=97235546; PubMed=9080665;  
RA Steiner J.M., Medinger T.L., Williams D.A.;  
RT "Purification and partial characterization of feline trypsin.";  
RL Comp. Biochem. Physiol. 116B:87-93(1997).  
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-  
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR InterPro; IPR001254; Trypsin.  
DR PROSITE; PS00240; TRYPSIN\_DOM; PARTIAL.  
DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPSIN\_SER; PARTIAL.  
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.  
FT PROPEP 1 8 ACTIVATION PEPTIDE.  
FT CHAIN 9 >16 TRYPSIN.  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1825 MW; A6D751BB58750A86 CRC64;

Query Match 19.5%; Score 24; DB 1; Length 16;  
Best Local Similarity 66.7%; Pred. No. 9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PIPDED 10  
||:|:|:  
Db 2 PIDDDD 7

## RESULT 5

ID COXA\_THUOB STANDARD; PRT; 20 AA.  
AC P80972;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-1 (EC 1.9.3.1) (FRAGMENT).  
OS Thunnus obesus (Bigeye tuna).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;  
OC Scombridae; Thunnus.  
OX NCBI\_TaxID=8241;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Heart, and Liver;  
RX MEDLINE=97454291; PubMed=9310366;  
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,  
RA Kadenbach B.;  
RT "The subunit structure of cytochrome-c oxidase from tuna heart and

```

RT liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -|- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC -|- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -|- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2404 MW; 7E82E43B7157355E CRC64;

Query Match 19.5%; Score 24; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 DSPWIT 15
DB 12 DARWVT 17

RESULT 6
LMT2_L0CMI STANDARD; PRT; 8 AA.
ID LMT2_L0CMI
AC P22396;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE LOCUSTAMYOTROPIN 2 (LOM-MT-2).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT "Isolation, identification and synthesis of locustamyotropin II, an
RT additional neuropeptide of Locusta migratoria. Member of the
RT cephalomyotropic peptide family.";
RL Insect Biochem. 20:479-484(1990).
CC -|- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -|- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro; IPR001484; Pyrokinin.
DR PROSITE; PS00539; PYROKININ; 1.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8
FT AMIDATION.
SQ SEQUENCE 8 AA; 934 MW; 26341771A9CAA87B CRC64;

Query Match 18.7%; Score 23; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 DSTPRI 21
DB 3 DFTPLR 8

RESULT 7
PPAC_BACME STANDARD; PRT; 16 AA.
ID PPAC_BACME
AC P56948;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MANGANESE-DEPENDENT INORGANIC PYROPHOSPHATASE (EC 3.6.1.1)
DE (PYROPHOSPHATE PHOSPHO-HYDROLASE) (PPASE) (FRAGMENT).
GN PPAC.
OS Bacillus megaterium.

```

```

OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE.
RX MEDLINE=98455825; PubMed=9782505;
RA Young T.W., Kuhn N.J., Wadeson A., Ward S., Burges D., Cooke G.D.;
RT "Bacillus subtilis ORF yybQ encodes a manganese-dependent inorganic
RT pyrophosphatase with distinctive properties: the first of a new class
RT of soluble pyrophosphatase?";
RL Microbiology 144:2563-2571(1998).
CC -|- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O -> 2 ORTHOPHOSPHATE.
CC -|- COFACTOR: REQUIRES MANGANESE FOR ITS ACTIVITY.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE PPASE CLASS C FAMILY.
KW Hydrolase; Manganese.
FT NON_TER 16
FT SEQUENCE 16 AA; 1828 MW; 3C0E6735D98B38A0 CRC64;

Query Match 18.7%; Score 23; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 HPIPEDS 11
DB 8 HKNPDTDT 15

RESULT 8
LPK_LEUMA STANDARD; PRT; 8 AA.
ID LPK_LEUMA
AC P13049;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE LEUKOPYROKININ (LPK) (LEM-PK).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidae; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RX MEDLINE=86269041; PubMed=3015140;
RA Nachman R.J., Holman G.M., Cook B.J.;
RT "Active fragments and analogs of the insect neuropeptide
RT leucopyrokinin: structure-function studies.";
RL Biochem. Biophys. Res. Commun. 137:936-942(1986).
RN [2]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RX MEDLINE=87052651; PubMed=2877794;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Primary structure and synthesis of a blocked myotropic
RT neuropeptide isolated from the cockroach, Leucophaea maderae.";
RL Comp. Biochem. Physiol. 85C:219-224(1986).
CC -|- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -|- MISCELLANEOUS: AN ANALOG WITHOUT THE N-TERMINAL PCA RESIDUE WAS
CC SYNTHESIZED AND FOUND TO EXHIBIT GREATER ACTIVITY (144%) THAN THE
CC PARENT NEUROPEPTIDE. THE PORTION OF THE SEQUENCE OF LPK MOST
CC CRITICAL FOR THE MYOTROPIC PROPERTIES IS LIMITED TO THE
CC PENTAPEPTIDE FRAGMENT FTPLR.
CC -|- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro; IPR001484; Pyrokinin.
DR PROSITE; PS00539; PYROKININ; 1.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 1
FT AMIDATION.
FT MOD_RES 8
FT SEQUENCE 8 AA; 949 MW; 92341771A9D5A1B6 CRC64;
SQ SEQUENCE 8 AA; 949 MW; 92341771A9D5A1B6 CRC64;

```

```

Query Match          17.9%; Score 22; DB 1; Length 8;
Best Local Similarity 57.1%; Pred. No. le+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TDSTPRI 21
   1 |||
Db 2 TSFTPL 8

RESULT 9
PKCL_CARMO
ID PKCL_CARMO STANDARD; PRT; 11 AA.
AC PG2684;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PYROKININ-1 (CAM-PK-1) (FXPL-AMIDE).
OS Carausius morosus (Indian stick insect).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Phasmatodea; Heteronemiidae;
OC Carausius.
OX NCBI_TaxID=7022;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Corpora cardiaca;
RA Predel R., Kellner R., Gaede G.;
RT "Myotopic neuropeptides from the retrocerebral complex of the stick
insect, Carausius morosus (Phasmatodea: Lonchodidae).";
RL Eur. J. Entomol. 96:275-278(1999).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(NYOTROPIC ACTIVITY).
CC -1- MASS SPECTROMETRY: MW=1235; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro: IPR001484; Pyrokinin.
DR PROSITE: PS00539; PYROKININ; FALSE_NEG.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 11 11 AMIDATION.
SQ SEQUENCE 11 AA; 1236 MW; 2BFA5225BB46C1A8 CRC64;

```

```

Query Match          17.9%; Score 22; DB 1; Length 11;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TDSTPRI 21
   |   | | |
Db  5 TQYTPRL 11

RESULT 10
TKN1_PSEGU
ID   TKN1_PSEGU      STANDARD;          PRT;      11 AA.
AC  P42986;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  KASSININ-LIKE PEPTIDE K-1 (PG-K1).
OS  Pseudophryne guentheri (Frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC  Pseudophryne.
OX  NCBI_TaxID=30349;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Skin;
RX  MEDLINE=90287814; PubMed=2356157;
RA  Sirmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA  Roberts J.D., Welchiorri P., Ersamer V.;
RT  "Six novel tachykinin- and bombesin-related peptides from the skin of
RL  the Australian frog Pseudophryne guentheri.";
RL  Peptides 11:299-304(1990).
CC  -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS.

```

CC	EVOKED BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC	SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC	MUSCLES.
CC	- - SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC	PIR: B60409; B60409.
DR	InterPro: IPR003580; Protachykinin.
DR	InterPro: IPR002040; Tachykinin.
DR	Pfam: PF02202; Tachykinin; 1.
DR	SMART: SM00203; TK; 1.
DR	PROSITE: PS00267; TACHYKININ; 1.
KW	Tachykinin; Neuropeptide; Amidation.
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT	MOD_RES 11 11 AMIDATION.
SQ	SEQUENCE 11 AA; 1269 MW; 3DBA7C37C9CBIAB7 CRC64;
Query Match	17.9%; Score 22; DB 1; Length 11;
Best Local Similarity	80.0%; Pred. No. 1.2e+03;
Matches	4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	5 PIPDE 9
	1
Db	2 PHPDE 6
RESULT 11	
TKN2_PSEGU	STANDARD; PRT; 11 AA.
ID TKN2_PSEGU	
AC P42987;	
DT 01-NOV-1995 (Rel. 32, Created)	
DT 01-NOV-1995 (Rel. 32, Last sequence update)	
DT 30-MAY-2000 (Rel. 39, Last annotation update)	
DE KASSININ-LIKE PEPTIDE K-II (PG-KII).	
OS Pseudophryne guntheri (Frog).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;	
OC Pseudophryne.	
OX NCBI_TaxID=30349;	
XP [1]	
RN SEQUENCE.	
RC TISSUE=Skin;	
RX MEDLINE=90287814; PubMed=2356157;	
RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,	
RA Roberts J.D., Melchiorri P., Erspamer V.;	
RT "Six novel tachykinin- and bombesin-related peptides from the skin of	
RT the Australian frog Pseudophryne guntheri.";	
RL Peptides 11:299-304(1990).	
CC - - FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,	
CC EVOKED BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND	
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH	
CC MUSCLES.	
CC - - SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.	
CC PIR: C60409; C60409.	
DR InterPro: IPR003580; Protachykinin.	
DR InterPro: IPR002040; Tachykinin.	
DR Pfam: PF02202; Tachykinin; 1.	
DR SMART: SM00203; TK; 1.	
DR PROSITE: PS00267; TACHYKININ; 1.	
KW Tachykinin; Neuropeptide; Amidation.	
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT	MOD_RES 11 11 AMIDATION.
SQ	SEQUENCE 11 AA; 1246 MW; 3A247C37C9CBIAB7 CRC64;
Query Match	17.9%; Score 22; DB 1; Length 11;
Best Local Similarity	80.0%; Pred. No. 1.2e+03;
Matches	4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	5 PIPDE 9
	1
Db	2 PNPD 6

```

RESULT 12
TKNA_PSEGU STANDARD; PRT; 11 AA.
AC P42989;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SUBSTANCE P-LIKE PEPTIDE I (PG-SPI).
OS Pseudophryne guentheri (Frog).
OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC Pseudophryne.
OC NCBI_TaxID=30349;
RN [1]
RP SEQUENCE.
RX TISSUE=Skin;
RX MEDLINE=90287814; PubMed=2356157;
RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA Roberts J.D., Melchiorri P., Erspamer V.;
RT "Six novel tachykinin- and bombesin-related peptides from the skin of
RT the Australian frog Pseudophryne guntheri.";
RL Peptides 11:299-304(1990).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC PIR; E60409; E60409.
DR InterPro; IPR003580; Protachykinin.
DR InterPro; IPR002040; Tachykinin.
DR Pfam; PF02202; Tachykinin; 1.
DR SMART; SM00203; TK; 1.
DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 11 11 AMIDATION.
SQ SEQUENCE 11 AA; 1294 MW; 3A247C2CC9CB1AB7 CRC64;

Query Match 17.9%; Score 22; DB 1; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PIPDE 9
DB 2 PNPDE 6

RESULT 13
SODM_CANFA STANDARD; PRT; 13 AA.
AC P54712;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
GN SOD2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
```

```

CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC HSSP; P04179; IMSD.
DR HSC-2DPAGE; P54712; DOG.
DR InterPro; IPR001189; SOD_MI.
DR Pfam; PF00081; sodfe; 1.
DR PROSITE; PS00088; SOD_MN; PARTIAL.
KW Oxidoreductase; Manganese; Mitochondrion.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1476 MW; 9C9651DE8BE0672A CRC64;

Query Match 17.9%; Score 22; DB 1; Length 13;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 HPIPD 8
DB 2 HSLPD 6

RESULT 14
FIBB_PIG STANDARD; PRT; 19 AA.
AC P14477;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last annotation update)
DE FIBRINOPEPTIDE B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT MOD_RES 4 4 SULFATION.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2200 MW; 09F87E44F4F3863D CRC64;

Query Match 17.9%; Score 22; DB 1; Length 19;
Best Local Similarity 35.7%; Pred. No. 2.1e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 4; Gaps 1;

QY 8 DEDPWITDTPRI 21
DB 5 DEDE---DGRPKV 14

RESULT 15
ECDI_LYMDI STANDARD; PRT; 21 AA.
AC P80936; P55898;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TESTIS ECDYSIOTROPIN PEPTIDE 1 (TE).
```

OS Lymantria dispar (Gypsy moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Noctuoidea; Lymantriidae; Lymantria.  
 OX NCBI\_TaxID=13123;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RA Wagner R.M., Loeb M.J., Kochansky J.P., Gelman D.B., Lusby W.R.,  
 RA Bell R.A.;  
 RT "Identification and characterization of an ecdysiotropic peptide from  
 brain extracts of the the gypsy moth, Lymantria dispar.";  
 RL Arch. Insect Biochem. Physiol. 34:175-189(1997).  
 CC -!- FUNCTION: STIMULATES OR BOOSTS ECDYSTEROID SYNTHESIS IN TESTIS OF LARVAE  
 AND PUPAE.  
 SQ SEQUENCE 21 AA; 2474 MW; E6E819624D1DEB62 CRC64;

Query Match 17.9%; Score 22; DB 1; Length 21;  
 Best Local Similarity 33.3%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 PIPDESPWITD 16  
 | : | | : : |  
 Db 9 PLNDADNNEVLID 20

Search completed: February 21, 2002, 16:52:21  
 Job time: 520 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:51:54 ; Search time 60.03 Seconds

(without alignments)  
51.170 Million cell updates/sec

Title: US-08-753-851-6

Perfect score: 123

Sequence: 1 CTVHPIDEDSPWTDSTPRI 21

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 5819

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTREMBL\_17.\*

2: sp\_archaea.\*

3: sp\_bacteria.\*

4: sp\_fungi.\*

5: sp\_human.\*

6: sp\_invertebrate.\*

7: sp\_mammal.\*

8: sp\_mmc.\*

9: sp\_organalle.\*

10: sp\_phase.\*

11: sp\_plant.\*

12: sp\_rodent.\*

13: sp\_virus.\*

14: sp\_vertebrate.\*

15: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	25.2	16	Q9UC99	Q9UC99 homo sapien
2	31	25.2	20	Q61871	Q61871 mus musculu
3	30	24.4	15	Q40562	Q40562 nicotiana t
4	30	24.4	15	Q40563	Q40563 nicotiana t
5	30	24.4	18	Q62256	Q62256 mus musculu
6	30	24.4	20	Q9QU24	Q9QU24 rattus sp.
7	30	24.4	21	Q9X3L7	Q9X3L7 prochloroco
8	29	23.6	15	Q9QU23	Q9QU23 rattus sp.
9	27	22.0	10	Q9QVF7	Q9QVF7 rattus sp.
10	27	22.0	17	Q9UEJ5	Q9UEJ5 homo sapien
11	27	22.0	19	Q31687	Q31687 artemia par
12	27	22.0	19	Q9YR03	Q9YR03 porcine cir
13	26	21.1	10	Q9N1X1	Q9N1X1 equus cabal
14	26	21.1	18	Q9UGN8	Q9UGN8 homo sapien
15	26	21.1	20	P82258	P82258 aspergillus
16	26	21.1	20	Q9TRI3	Q9TRI3 bos taurus
17	26	21.1	21	Q83489	Q83489 turnip yell
18	25.5	20.7	20	Q9TR87	Q9TR87 bos taurus
19	25	20.3	14	Q9TWX8	Q9TWX8 manduca sex

20	25	20.3	17	4	Q9HAR9	Q9HAR9 homo sapien
21	25	20.3	18	4	Q9Y6D8	Q9Y6D8 homo sapien
22	25	20.3	18	11	P70649	P70649 mus musculu
23	25	20.3	19	5	Q9TWH2	Q9TWH2 oxytricha g
24	25	20.3	19	12	O91329	O91329 human immun
25	25	20.3	20	3	Q9URC7	Q9URC7 saccharomyc
26	25	20.3	20	6	Q9N188	Q9N188 equus cabal
27	25	20.3	21	13	Q9PS51	Q9PS51 lampetra fl
28	24.5	19.9	18	10	Q9S915	Q9S915 triticum tu
29	24.5	19.9	18	10	Q9S817	Q9S817 oryza sativ
30	24	19.5	12	2	Q46747	Q46747 escherichia
31	24	19.5	14	11	Q9QVD9	Q9QVD9 rattus sp.
32	24	19.5	15	2	Q9R599	Q9R599 micrococcu
33	24	19.5	19	2	Q45558	Q45558 bacillus sp
34	24	19.5	19	2	Q9EV48	Q9EV48 staphylococ
35	24	19.5	19	11	Q9QV70	Q9QV70 rattus sp.
36	24	19.5	19	12	Q9YIT3	Q9YIT3 bovine circ
37	24	19.5	19	12	O56132	O56132 porcine cir
38	24	19.5	20	2	Q9R4H4	Q9R4H4 desulfovibr
39	24	19.5	21	11	Q9QV03	Q9QV03 rattus sp.
40	23.5	19.1	17	4	Q9UDD6	Q9UDD6 homo sapien
41	23	18.7	15	10	Q9S929	Q9S929 glycine max
42	23	18.7	17	6	O62645	O62645 saquinavir
43	23	18.7	19	13	Q9PRN1	Q9PRN1 petromyzon
44	23	18.7	20	11	Q62546	Q62546 mus spretus
45	23	18.7	21	2	Q9X3G6	Q9X3G6 prochloroco

## ALIGNMENTS

RESULT 1

Q9UC99 PRELIMINARY; PRT; 16 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE EOCIP-1-EOSINOPHIL-CHEMOTACTIC CYTOKINE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=95028770; PubMed=7524281;  
RA Schroder J.M., Kameyoshi Y., Christophers E.;

RT "Platelets secrete an eosinophil-chemotactic cytokine which is a member of the C-C-chemokine family.";  
RL Adv. Exp. Med. Biol. 351:119-128(1993).  
SQ SEQUENCE 16 AA; 1786 MW; 8B041B46AEACC2A1 CRC64;

Query Match 25.2%; Score 31; DB 4; Length 16;  
Best Local Similarity 55.6%; Pred. No. 3.6e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 SPWITDSTP 19

Db 1 SPYXXDTTP 9

RESULT 2

Q61871 PRELIMINARY; PRT; 20 AA.

AC Q61871;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE ENVELOPE PROTEIN (FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AKR; TISSUE=THYMUS;  
 RX MEDLINE=88091045; PubMed=2826802;  
 RA Laigret F., Repaske R., Boulukos K., Rabson A.B., Khan A.S.;  
 RT "Potential progenitor sequences of mink cell focus-forming (MCF)  
 murine leukemia viruses: ecotropic, xenotropic, and MCF-related viral  
 RNAs are detected concurrently in thymus tissues of AKR mice.";  
 RL J. Virol. 62:376-386(1988).  
 DR EMBL: M19049; AAA18408.1; -  
 DR InterPro: IPR002050; Env\_polyprotein.  
 DR Pfam: PF00429; Env\_polyprotein; 1.  
 KW Envelope protein.  
 FT NON\_TER 20  
 SQ SEQUENCE 20 AA; 2240 MW; 285338BE2EB5C9DD CRC64;

Query Match 25.2%; Score 31; DB 11; Length 20;  
 Best Local Similarity 44.4%; Pred. No. 4.6e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 PIPDESPW 13  
 Db 9 PLKDKINPW 17

RESULT 3  
 Q40562 PRELIMINARY; PRT; 15 AA.  
 ID Q40562;  
 AC Q40562;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE SMALL RAS-RELATED PROTEIN (FRAGMENT).  
 GN RAN-A2.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SRI;  
 RX MEDLINE=95078947; PubMed=7987414;  
 RA Merkle T., Haizel T., Matsumoto T., Harter K., Dallmann G., Nagy F.;  
 RT "Phenotype of the fission yeast cell cycle regulatory mutant pim1-46  
 is suppressed by a tobacco cDNA encoding a small, Ran-like GTP-binding  
 protein.";  
 RL Plant J. 6:555-565(1994).  
 DR EMBL: L16786; AAA73564.1; -  
 KW Cell cycle; GTP-binding.  
 FT NON\_TER 1  
 SQ SEQUENCE 15 AA; 1707 MW; 8AB89F924EA30CAF CRC64;

Query Match 24.4%; Score 30; DB 10; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 PIPDED 10  
 Db 4 PLPDDD 9

RESULT 4  
 Q40563 PRELIMINARY; PRT; 15 AA.  
 ID Q40563;  
 AC Q40563;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE SMALL RAS-RELATED PROTEIN (FRAGMENT).

GN RAN-B2.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SRI;  
 RX MEDLINE=95078947; PubMed=7987414;  
 RA Merkle T., Haizel T., Matsumoto T., Harter K., Dallmann G., Nagy F.;  
 RT "Phenotype of the fission yeast cell cycle regulatory mutant pim1-46  
 is suppressed by a tobacco cDNA encoding a small, Ran-like GTP-binding  
 protein.";  
 RL Plant J. 6:555-565(1994).  
 DR EMBL: L16788; AAA73565.1; -  
 KW Cell cycle; GTP-binding.  
 FT NON\_TER 1  
 SQ SEQUENCE 15 AA; 1604 MW; 9A57F48F1FFFFFAF CRC64;

Query Match 24.4%; Score 30; DB 10; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 PIPDED 10  
 Db 5 PLPDDD 10

RESULT 5  
 Q62256 PRELIMINARY; PRT; 18 AA.  
 ID Q62256;  
 AC Q62256;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE SPERMATOGENIC-SPECIFIC PROENKEPHALIN.  
 GN PENK2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90287163; PubMed=2355920;  
 RA Kilpatrick D.L., Zinn S.A., Fitzgerald M., Higuchi H., Sabol S.L.,  
 RA Meyerhardt J.;  
 RT "Transcription of the rat and mouse proenkephalin genes is initiated  
 at distinct sites in spermatogenic and somatic cells.";  
 RL Mol. Cell. Biol. 10:3717-3726(1990).  
 DR EMBL: M55181; AAA40127.1; -  
 DR MGD; MGI:104628; Penk2.  
 SQ SEQUENCE 18 AA; 2043 MW; B96E10CC7049FA76 CRC64;

Query Match 24.4%; Score 30; DB 11; Length 18;  
 Best Local Similarity 80.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 EDSPW 13  
 Db 6 QDSPW 10

RESULT 6  
 Q9QUZ4 PRELIMINARY; PRT; 20 AA.  
 ID Q9QUZ4;  
 AC Q9QUZ4;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE BETA-CRYSTALLIN ISOFORM A3 (FRAGMENT).

OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96030616; PubMed=7556477;  
 RA Shearer T.R., Shih M., Azuma M., David L.L.;  
 RT "Precipitation of crystallins from young rat lens by endogenous  
 RT calpain.";  
 RL Exp. Eye Res. 61:141-150(1995).  
 SQ SEQUENCE 20 AA; 2147 MW; C5860A30C6220681 CRC64;

Query Match 24.4%; Score 30; DB 11; Length 20;  
 Best Local Similarity 40.0%; Pred. No. 6.4e+02;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 HPIPEDSPW 13  
 :|:| |  
 Db 11 NPMPSMGWP 20

RESULT 7

QX3L7 ID Q9X3L7 PRELIMINARY; PRT; 21 AA.  
 AC Q9X3L7;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).  
 GN PETD.  
 OS Prochlorococcus sp.  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=1220;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Urbach E., Chisholm S.W.;  
 RT "Genetic diversity in Prochlorococcus populations flow cytometrically  
 RT sorted from the Sargasso Sea and Gulf Stream.";  
 RL Limnol. Oceanog. 43:1615-1630(1998).  
 DR EMBL; AF070208; AAD23258.1; -.  
 FT NON\_TER 21  
 SQ SEQUENCE 21 AA; 2360 MW; CC83509577BFF27F CRC64;

Query Match 24.4%; Score 30; DB 2; Length 21;  
 Best Local Similarity 42.9%; Pred. No. 6.8e+02;  
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 PDEDSWITDTPR 20  
 || | : | : |  
 Db 7 PDLADPKLKQSSPR 20

RESULT 8

Q9QUZ3 ID Q9QUZ3 PRELIMINARY; PRT; 15 AA.  
 AC Q9QUZ3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE BETA-CRYSTALLIN ISOFORM A3/A1 (FRAGMENT).  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96030616; PubMed=7556477;  
 RA Shearer T.R., Shih M., Azuma M., David L.L.;  
 RT "Precipitation of crystallins from young rat lens by endogenous

RT calpain.";  
 RL Exp. Eye Res. 61:141-150(1995).  
 SQ SEQUENCE 15 AA; 1702 MW; 3F35688E1C5F233C CRC64;

Query Match 23.6%; Score 29; DB 11; Length 15;  
 Best Local Similarity 44.4%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 PIPEDSPW 13  
 |:| |  
 Db 2 PMPSMGWP 10

RESULT 9

Q9QVF7 ID Q9QVF7 PRELIMINARY; PRT; 10 AA.  
 AC Q9QVF7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE BETA 2-GLYCOPROTEIN I, BETA 2-GPI.  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92135065; PubMed=1777418;  
 RA Matsuura E., Igarashi M., Igarashi Y., Nagae H., Ichikawa K.,  
 RA Yasuura T., Koike T.;  
 RT "Molecular definition of human beta 2-glycoprotein I (beta 2-GPI) by  
 RT cDNA cloning and inter-species differences of beta 2-GPI in  
 RT alternation of antidiolipin binding.";  
 RL Int. Immunol. 3:1217-1221(1991)  
 SQ SEQUENCE 10 AA; 1100 MW; 94E681B767376E1 CRC64;

Query Match 22.0%; Score 27; DB 11; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 8.3e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 PIPDE 9  
 |:| |  
 Db 5 PLPDE 9

RESULT 10

Q9UEJ5 ID Q9UEJ5 PRELIMINARY; PRT; 17 AA.  
 AC Q9UEJ5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE PLAU PROTEIN (FRAGMENT).  
 GN PLAU.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86050639; PubMed=3933505;  
 RA Nagamine Y., Pearson D., Grattan M.;  
 RT "Exon-intron boundary sliding in the generation of two mRNAs coding  
 RT for porcine urokinase-like plasminogen activator.";  
 RL Biochem. Biophys. Res. Commun. 132:563-569(1985).  
 DR EMBL; K03027; AAA61257.1; -.  
 DR HSSP; P00749; IURK.  
 FT NON\_TER 1  
 FT NON\_TER 17  
 SQ SEQUENCE 17 AA; 1801 MW; 9989927EA849E253 CRC64;

Query Match 22.0%; Score 27; DB 4; Length 17;  
Best Local Similarity 41.7%; Pred. No. 1.5e+03;  
Matches 5; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTVHPIDPDESP 12  
I I I I I  
DB 2 CMVHDCADGKKP 13

## RESULT 11

ID Q31687 PRELIMINARY; PRT; 19 AA.  
AC Q31687;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE ATPASE SUBUNIT 8 (FRAGMENT).  
GN ATP8.  
OS Artemia parthenogenetica.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;  
OC Artemiidae; Artemia.  
OX NCBI\_TaxID=6663;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LA MATA.  
RX MEDLINE=94223692; PubMed=8159960;  
RA Perez M.L., Valverde J.R., Batuecas B., Amat F., Marco R., Garesse R.;  
RT "Speciation in the Artemia genus: mitochondrial DNA analysis of  
RT bisexual and parthenogenetic brine shrimps."  
RL J. Mol. Evol. 38:156-168(1994).  
DR EMBL: X67263; CAA47685.1; -.  
KW Mitochondrion.  
FT NON\_TER 1  
FT NON\_TER 19  
SQ SEQUENCE 19 AA; 2246 MW; 41922AD313E087E3 CRC64;

Query Match 22.0%; Score 27; DB 8; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PWIT 15  
I I I I I  
DB 8 PWIT 11

## RESULT 12

ID Q9YR03 PRELIMINARY; PRT; 19 AA.  
AC Q9YR03;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE ORF-7.  
OS porcine circovirus type 2-E.  
OC Viruses; ssDNA viruses; Circoviridae; Circovirus.  
OX NCBI\_TaxID=85544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hamel A.L., Nayar G.P.S.;  
RT "Nucleotide sequence of four different isolates of circovirus detected  
RT in pigs with various clinical syndromes."  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF109399; AAD03077.1; -.  
SQ SEQUENCE 19 AA; 1955 MW; 2FD0BE58EAF9A923 CRC64;

Query Match 22.0%; Score 27; DB 12; Length 19;  
Best Local Similarity 55.6%; Pred. No. 1.7e+03;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 EDSPWITDS 17  
I I I I I  
DB 11 ETPPWIRHS 19

## RESULT 13

ID Q9N1X1 PRELIMINARY; PRT; 10 AA.  
AC Q9N1X1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE ALCOHOL DEHYDROGENASE 3 (FRAGMENT).  
GN ADH3.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20082971; PubMed=10613847;  
RA Caetano A.R., Shlue Y.L., Lyons L.A., O'Brien S.J., Laughlin T.F.,  
RA Bowling A.T., Murray J.D.;  
RT "A comparative gene map of the horse (Equus caballus).";  
RT Genome Res. 9:1239-1249(1999).  
DR EMBL: AF134056; AAF31299.1; -.  
FT NON\_TER 1  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1144 MW; C4EA25676B02D6DD CRC64;

Query Match 21.1%; Score 26; DB 6; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 PWIT 15  
I I I I I  
DB 7 PWIT 10

## RESULT 14

ID Q9UGN8 PRELIMINARY; PRT; 18 AA.  
AC Q9UGN8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE APAF1 PROTEIN (FRAGMENT).  
GN APAF1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Roberts D.L., Dalglish R., Cohen G.M., MacFarlane M.;  
RT "The mammalian CED4 homologue, APAF1, exists as two distinct forms in  
RT human cells."  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ133645; CAB65087.1; -.  
FT NON\_TER 1  
FT NON\_TER 18  
SQ SEQUENCE 18 AA; 2045 MW; 30D5FA30B885AEF5 CRC64;

Query Match 21.1%; Score 26; DB 4; Length 18;  
Best Local Similarity 44.4%; Pred. No. 2.2e+03;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 13 WITDSTPRI 21  
I I I I I  
DB 2 WNTDSRSKV 10

RESULT 15  
P82258  
ID P82258 PRELIMINARY; PRT; 20 AA.  
AC P82258;  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)  
DE ALLERGEN ASP L 2 (FRAGMENT).  
OS Aspergillus flavus.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
OX NCBI\_TaxID=5059;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=AFL-1505 / INDIAN ISOLATE;  
RA Sarma P.U., Paliwal A., Fairwell T.;  
RL Submitted (DEC-1999) to the SWISS-PROT data bank.  
CC -!- FUNCTION: BINDS TO IGG AND IGE.  
KW Allergen; IgE-binding protein; IgG-binding protein.  
FT NON\_TER 1  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2446 MW; 40C642784C2A4A13 CRC64;

Query Match 21.1%; Score 26; DB 3; Length 20;  
Best Local Similarity 50.0%; Pred. No. 2.5e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 8 DEDSPW 13  
Db 2 DRNPW 7

Search completed: February 21, 2002, 16:51:54  
Job time: 533 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:42:57 ; Search time 64.88 Seconds  
(without alignments)  
23.976 Million cell updates/sec

Title: US-08-753-851-6

Perfect score: 123

Sequence: 1 CTVHPIDEDSPWITDTPRI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 206698

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Genesec\_1101.\*

```

1: /SID88/gcgdata/genesec/genesecp/AA1980.DAT.*
2: /SID88/gcgdata/genesec/genesecp/AA1981.DAT.*
3: /SID88/gcgdata/genesec/genesecp/AA1982.DAT.*
4: /SID88/gcgdata/genesec/genesecp/AA1983.DAT.*
5: /SID88/gcgdata/genesec/genesecp/AA1984.DAT.*
6: /SID88/gcgdata/genesec/genesecp/AA1985.DAT.*
7: /SID88/gcgdata/genesec/genesecp/AA1986.DAT.*
8: /SID88/gcgdata/genesec/genesecp/AA1987.DAT.*
9: /SID88/gcgdata/genesec/genesecp/AA1988.DAT.*
10: /SID88/gcgdata/genesec/genesecp/AA1989.DAT.*
11: /SID88/gcgdata/genesec/genesecp/AA1990.DAT.*
12: /SID88/gcgdata/genesec/genesecp/AA1991.DAT.*
13: /SID88/gcgdata/genesec/genesecp/AA1992.DAT.*
14: /SID88/gcgdata/genesec/genesecp/AA1993.DAT.*
15: /SID88/gcgdata/genesec/genesecp/AA1994.DAT.*
16: /SID88/gcgdata/genesec/genesecp/AA1995.DAT.*
17: /SID88/gcgdata/genesec/genesecp/AA1996.DAT.*
18: /SID88/gcgdata/genesec/genesecp/AA1997.DAT.*
19: /SID88/gcgdata/genesec/genesecp/AA1998.DAT.*
20: /SID88/gcgdata/genesec/genesecp/AA1999.DAT.*
21: /SID88/gcgdata/genesec/genesecp/AA2000.DAT.*
22: /SID88/gcgdata/genesec/genesecp/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	123	100.0	21	CD44 peptide CD44-
2	40	32.5	20	HSV truncated tegu
3	39.5	32.1	18	Peptide #2092 enco
4	39.5	32.1	22	Peptide #2196 enco
5	39.5	32.1	18	Peptide #2077 enco
6	36	29.3	17	Peptide #5652 enco
7	36	29.3	17	Peptide #5912 enco
8	34	27.6	14	Peptide #7 inserte
9	34	27.6	14	Randomly generated
10	34	27.6	14	Thioredoxin active
11	34	27.6	18	Human bcl2 proto-o

```

12 34 27.6 18 19 AAW66667
13 34 27.6 20 21 AAY67265
14 33 26.8 14 18 AAW4962
15 33 26.8 16 18 AAW25373
16 33 26.8 20 11 AAR04988
17 33 26.8 20 17 AAW17007
18 33 26.8 20 21 AAY67260
19 33 26.8 20 21 AAY67267
20 33 26.8 21 10 AAY91156
21 33 26.8 21 20 AAY17230
22 33 26.8 21 20 AAY17232
23 32.5 26.4 17 21 AAB34527
24 32 26.0 12 22 AAB65576
25 32 26.0 16 18 AAW25402
26 32 26.0 19 15 AAR60808
27 32 26.0 20 20 AAY38490
28 32 26.0 20 21 AAY67264
29 32 26.0 21 15 AAR60803
30 31.5 25.6 15 16 AAR72634
31 31.5 25.6 17 16 AAR72655
32 31 25.2 9 21 AAY80169
33 31 25.2 13 20 AAY06588
34 31 25.2 20 9 AAP80172
35 31 25.2 20 22 AAB55793
36 31 25.2 20 22 AAB57655
37 31 25.2 20 22 AAB56065
38 31 25.2 21 18 AAW16218
39 31 25.2 21 19 AAW68851
40 31 25.2 21 19 AAW38339
41 31 25.2 21 20 AAY09987
42 31 25.2 21 21 AAB17767
43 30.5 24.8 18 16 AAR79523
44 30.5 24.8 19 21 AAB34528
45 30.5 24.8 20 14 AAR33487

```

#### ALIGNMENTS

```

RESULT 1
AAR53478
ID AAR53478 standard; peptide; 21 AA.
XX
XX AAR53478;
XX
XX 01-DEC-1994 (first entry)
XX
XX CD44 peptide CD44-6.
XX
XX Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;
XX arthritis; tumor cell metastasis; autoimmune disease;
XX immunosuppressive.
XX
XX Homo sapiens.
XX
XX WO9409811-A.
XX
XX 11-MAY-1994.
XX
XX 29-OCT-1993; 93WO-US10412.
XX
XX 30-OCT-1992; 92US-0973339.
XX
XX (UYDU-) UNIV DUKE.
XX
XX Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;
XX WPI; 1994-167121/20.
XX
XX Use of CD44 protein and new peptide derivs - for developing prods
XX for inflammation, immune-mediated tissue damage and tumour cell
XX metastasis
XX

```

```

HSV-2 glycoprotein
Protein tyrosine p
C3/C4/C5-derived c
Src SH3 domain bin
Papilloma virus ty
SRC SH3 domain-bin
Protein tyrosine p
Protein tyrosine p
Artificial peptide
Casein antimicrobi
Casein antimicrobi
Human secreted pro
Immunogenic peptid
p53bp2 SH3 domain
Aspergillus niger
Human secreted pro
Protein tyrosine p
Aspergillus niger
Cladosporium herba
Cladosporium herba
HUA-A2 restricted
Neutral invertase
Sequence of P85 po
PL peptide #4. Sy
CD3n PL peptide.
CD3n PL peptide.
Peptide containing
Peptide binding in
IL-1RI binding pe
Interleukin-1 type
IL-1 antagonist pe
Hepatitis B virus
Human secreted pro
Cytotoxic T-lympho

```

AC AAM28159;

DT 17-OCT-2001 (first entry)  
DE Peptide #2196 encoded by probe for measuring placental gene expression.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00663.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488897/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX Claim 27; SEQ ID No 28428; 654pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENP:  
CC see AA131315-AA157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 18 AA;  
  
Query Match 32.1%; Score 39.5; DB 22; Length 18;  
Best Local Similarity 42.1%; Pred. No. 41;  
Matches 8; Conservative 2; Mismatches 4; Indels 5; Gaps 1;  
  
Qy 1 CTVHPIDEDSPWITDSTP 19  
II : III : III  
Db 5 ctt-----ksspcwcpnstp 18  
  
RESULT 5  
AAM03395  
ID AAM03395 standard; Protein; 18 AA.  
XX  
XX AAM03395;  
AC  
XX  
XX 09-OCT-2001 (first entry)  
XX  
XX Peptide #2077 encoded by probe for measuring breast gene expression.  
DE  
DE Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
KW  
XX Homo sapiens.  
OS  
XX WO200157270-A2.  
PN  
XX 09-AUG-2001.  
XX  
XX 29-JAN-2001; 2001WO-US00661.  
PF

XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-476286/51.  
XX  
XX Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
XX Claim 27; SEQ ID No 12135; 322pp; English.  
XX  
XX The present invention relates to novel single exon nucleic acid probes  
CC (see AA100010-AA110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 18 AA;  
  
Query Match 32.1%; Score 39.5; DB 22; Length 18;  
Best Local Similarity 42.1%; Pred. No. 41;  
Matches 8; Conservative 2; Mismatches 4; Indels 5; Gaps 1;  
  
Qy 1 CTVHPIDEDSPWITDSTP 19  
II : III : III  
Db 5 ctt-----ksspcwcpnstp 18  
  
RESULT 6  
AAM19218  
ID AAM19218 standard; Protein; 17 AA.  
XX  
XX AAM19218;  
AC  
XX  
XX 12-OCT-2001 (first entry)  
XX  
XX Peptide #5652 encoded by probe for measuring cervical gene expression.  
DE  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
KW  
XX Homo sapiens.  
OS  
XX WO200157278-A2.  
PN  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00670.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR



```

PR 04-OCT-2000; 2000GB-0024263.
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 24044; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 17 AA;

Query Match 29.3%; Score 36; DB 22; Length 17;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PDEDSPWIT 15
DQ | |||:|
Db 7 pssspwlt 15

RESULT 7
AAM31875
ID AAM31875 standard; Protein; 17 AA.
XX
AC AAM31875;
XX
DT 17-OCT-2001 (first entry)
DE Peptide #5912 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for

```

```

PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID No 32144; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 17 AA;

Query Match 29.3%; Score 36; DB 22; Length 17;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PDEDSPWIT 15
DQ | |||:|
Db 7 pssspwlt 15

RESULT 8
AAR26224
ID AAR26224 standard; Protein; 14 AA.
XX
AC AAR26224;
XX
DT 20-MAY-1998 (first entry)
DE Peptide #7 inserted into active site loop of thioredoxin.
XX
KW thioredoxin; soluble; stable; stability; research; diagnosis;
KW therapeutics; produce eukaryotic in prokaryotic; improved solubility;
KW enhanced heat stability; enhanced protease stability; ss.
XX
OS Homo sapiens.
XX
PN WO9213955-A.
XX
PD 20-AUG-1992.
XX
PF 06-FEB-1992; 92WO-US00944.
XX
PR 06-FEB-1991; 91US-0652531.
PR 14-AUG-1991; 91US-0745382.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Lavallie ER, McCoy J;
XX
DR WPI; 1992-300041/36.
XX
PT DNA sequence encoding a fusion protein - comprises thioredoxin
PT and a heterologous protein sequence; used to produce stable
PT soluble proteins e.g. IL-11
XX
PS Example 5; Page 31; 53pp; English.
XX
CC This sequence was used to make a fusion protein with E. coli thioredoxin
CC protein. The protein encodes an enterokinase cleavage site and was
CC inserted into the active site loop of thioredoxin between residues
CC G34 and P35. The fusion protein containing this internal
CC enterokinase site was expressed at levels equivalent to native
CC thioredoxin and was cleaved with an enterokinase treatment. The fusion
CC protein was found to be as stable as native thioredoxin, being stable to
CC a 10 minute incubation at 80C. It appears that the active-site loop of
CC thioredoxin can tolerate a wide variety of peptide insertions resulting
CC in soluble fusion proteins.
XX
SQ Sequence 14 AA;

```

```

Query Match      27.6%; Score 34; DB 13; Length 14;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 PWITDSTP 19
Db 1 pwingatp 8

RESULT 9
AAR45927
ID AAR45927 standard; peptide; 14 AA.
XX
AC AAR45927;
XX
DT 27-JUL-1994 (first entry)
XX
DE Randomly generated peptide 7 for insertion in E.coli thioredoxin.
XX
KW fusion gene; fusion protein; randomly generated peptide;
KW bacterial thioredoxin; E.coli; trxA gene; active site loop;
KW stability; increase; heterologous protein production;
KW protein folding.
XX
OS Synthetic.
XX
PN W09402502-A.
XX
PD 03-FEB-1994.
XX
PF 23-JUL-1993; 93WO-US06913.
XX
PR 28-JUL-1992; 92US-0921848.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Lavallie ER, McCoy J;
XX
DR WPI; 1994-048787/06.
XX
PT Recombinant prodn. of heterologous proteins - by expression as
PT fusion proteins with a thioredoxin-like protein for high
PT stability and solubility
XX
PS Claim 3 and Example 12; Page 32; 53pp; English.
XX
CC Twelve randomly generated peptide insertions were made into the
CC active site loop of E.coli thioredoxin, between residues Gly34 and
CC Pro35. Each of the thioredoxin fusion proteins containing these
CC random insertions of 14 amino acids were made at levels comparable
CC to native thioredoxin. All of them were found in the soluble
CC cellular fraction. The inserted sequences (AAR45921-R45932) include
CC examples that were both hydrophobic and hydrophilic, and examples
CC which contained Cys residues. It appears that the active site loop
CC of thioredoxin can tolerate a wide variety of peptide insertions
CC resulting in soluble fusion proteins.
XX
SQ Sequence 14 AA;

Query Match      27.6%; Score 34; DB 15; Length 14;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 PWITDSTP 19
Db 1 pwingatp 8

RESULT 10
AAR75776
ID AAR75776 standard; Peptide; 14 AA.

```

```

XX AAR75776;
XX
DT 20-NOV-1995 (first entry)
XX
DE Thioredoxin active site loop insert peptide.
XX
KW Thioredoxin; fusion protein.
XX
OS Synthetic.
XX
PN W09516044-A.
XX
PD 15-JUN-1995.
XX
PF 08-DEC-1994; 94WO-US14179.
XX
PR 10-DEC-1993; 93US-0165301.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Diblasio-Smith E, Grant K, Lavallie ER, McCoy J;
XX
DR WPI; 1995-224326/29.
XX
PT Proteins and peptide(s) fused to thioredoxin or thioredoxin-like
PT molecules - useful for prodn. of large amts. of heterologous proteins
XX
PS Claim 9; Page 39; 75pp; English.
XX
CC Small peptides (AAR75770-81) of random composition and hydrophilic,
CC hydrophobic or neutral character were inserted into the active site
CC loop (G34-P35) of E. coli thioredoxin (AAR75769). The resulting
CC fusion proteins were expressed in the soluble cellular fraction of
CC E. coli G1724 hosts at levels comparable to native thioredoxin.
XX
SQ Sequence 14 AA;

Query Match      27.6%; Score 34; DB 16; Length 14;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 PWITDSTP 19
Db 1 pwingatp 8

RESULT 11
AAY21170
ID AAY21170 standard; Protein; 18 AA.
XX
AC AAY21170;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human bcl2 proto-oncogene mutant protein fragment 18.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN W09845322-A2.

```

XX PD 15-OCT-1998.  
 XX PF 02-APR-1998; 98WO-1B00705.  
 XX PR 10-APR-1997; 97US-0043163.  
 XX PA (UYUT-) RIJKSUNIV UTRECHT.  
 XX PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
 XX PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
 XX PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
 XX WIPI; 1998-609901/51.  
 XX DR N-PSDB; AAX75766.  
 XX PT Diagnosing disease by detecting frameshift mutations in RNA or  
 XX PT corresponding protein mutations - used to diagnose cancer and  
 XX PT neurological diseases, particularly Alzheimer's disease, and also  
 XX PT for treatment and prevention with specific ribozymes or wild-type  
 XX PT RNA  
 XX PS Disclosure; Figure 15; 258pp; English.  
 XX CC This invention describes a novel method for the diagnosis of a disease  
 XX CC caused by, or associated with, an RNA molecule that has a frameshift  
 XX CC mutation. The method is used to diagnose age-related diseases, especially  
 XX CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
 XX CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
 XX CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
 XX CC and many others listed) or susceptibility to these disorders. The method  
 XX CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
 XX CC at an early stage. It is based on the observation that disease may be  
 XX CC caused by mutations in RNA rather than DNA. The invention describes the  
 XX CC use of neuronal system RNA molecules, specifically proteins including  
 XX CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
 XX CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
 XX CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
 XX CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic  
 XX CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
 XX CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
 XX CC protein-C (HMGP-C) and neuroendocrine specific protein A.  
 XX CC Sequence 18 AA;  
 XX QY 1 CTVHPIDEDSPWI 14  
 XX DB : : : : :  
 XX 3 cswwpwwelaspwv 16  
 XX  
 XX RESULT 12  
 XX AAW66667  
 XX ID AAW66667 standard; peptide; 18 AA.  
 XX AC AAW66667;  
 XX AC  
 XX DT 30-NOV-1998 (first entry)  
 XX DE HSV-2 glycoprotein G partial sequence (residues 451-468).  
 XX KW Herpes simplex virus type 2; HSV-2 antibody; detection; HSV-1;  
 XX KW multiply displayed peptide structure.  
 XX OS Synthetic.  
 XX OS Herpes simplex virus type 2.  
 XX PN GB2323360-A.  
 XX

XX PD 23-SEP-1998.  
 XX PF 14-JAN-1997; 97GB-0000660.  
 XX PR 14-JAN-1997; 97GB-0000660.  
 XX PA (MEDI-) MEDICAL RES COUNCIL.  
 XX PI Marsden HS;  
 XX WIPI; 1998-459516/40.  
 XX PT New peptide compounds useful for detecting herpes simplex virus type  
 XX PT 2 - can differentiate between herpes simplex virus type 1 and type 2.  
 XX PS Example 1; Page 10; 25pp; English.  
 XX CC The invention relates to a multiply displayed peptide structure of  
 XX CC formula [(X1)p-A-(X2)q-Sp]n-Core, X1, X2 = 1-6 non-interfering amino  
 XX CC acid residues; A = Glu Gly Phe Glu Ala Gly Asp Gly Glu Pro Glu  
 XX CC Asp Asp Asp; Sp = spacer group extending outwardly from the core; n at  
 XX CC least 4; and p, q = 0 or 1. The linkage between the core and the spacer  
 XX CC may be chemical or physical. Peptides of this formula are used in the  
 XX CC diagnosis of herpes simplex virus type 2 (HSV-2). The peptides can  
 XX CC distinguish HSV-2 from HSV-1. A series of 67 peptides (AAW66624-W66690),  
 XX CC mostly 18 amino acids long, that spanned amino acids 21-699 of the  
 XX CC predicted open reading frame of HSV gC2 were synthesised. The peptides  
 XX CC were made as multiply displayed peptide structures of the invention and  
 XX CC were screened against sera from HSV-1, HSV-2 antibody-positive  
 XX CC individuals and from individuals having no laboratory evidence of HSV  
 XX CC infection. From the results, peptide 55 (AAW66678) was considered a  
 XX CC likely candidate for type specific serodiagnosis of HSV.  
 XX CC Sequence 18 AA;  
 XX QY 5 PIPDESPWITDSTPR 20  
 XX DB : : : : :  
 XX 1 papptppptsthatpr 16  
 XX  
 XX RESULT 13  
 XX AAY67265  
 XX ID AAY67265 standard; protein; 20 AA.  
 XX AC AAY67265;  
 XX AC  
 XX DT 05-APR-2000 (first entry)  
 XX DE Protein tyrosine phosphatase (PTP) PEST P369A mutant peptide sequence.  
 XX KW Protein tyrosine phosphatase; PTP; PEST; paxillin binding; adhesion;  
 XX KW cell migration; division; cytostatic; antiinflammatory; angiogenesis;  
 XX KW cancer; enzyme substrate identification; muten.  
 XX OS Synthetic.  
 XX PH Key Location/Qualifiers  
 XX FT Misc-difference 15 /note= "wild type Pro been replaced by Ala"  
 XX FT  
 XX PN WO9961467-A2.  
 XX XX  
 XX PD 02-DEC-1999.  
 XX XX  
 XX XX 21-MAY-1999; 99WO-CA00461.  
 XX PF 21-MAY-1998; 98CA-2238654.  
 XX PR 11-DEC-1998; 98US-0111993.  
 XX

XX PA (UYMC-) UNIV MCGILL.  
 XX PI Tremblay ML, Cote J, Angers-Lousteau A, Charest A;  
 XX PI WPI; 2000-097104/08.  
 XX DR Novel therapeutic agents for treating diseases which are related to  
 PT cell proliferation, migration, inflammation and angiogenesis especially  
 PT cancer  
 XX PS Example 4; Fig 18; 91pp; English.  
 XX CC This is a fragment of the protein tyrosine phosphatase (PTP) PEST, P369A  
 CC mutant protein sequence. The fragment corresponds to amino acids  
 CC 355-374, and has the wild type proline residue corresponding to position  
 CC 369 replaced by an alanine residue. PTP-PEST is a soluble PTP that is  
 CC ubiquitously expressed throughout embryonic development and in murine  
 CC adult tissues. The N-terminal portion of the enzyme encodes for the  
 CC catalytic domain, while the C-terminal portion is composed of 5 proline  
 CC rich domains, and a binding site for the adaptor protein Shc. The pro 2  
 CC domain is required for paxillin binding, and the synthesis of mutant  
 CC PTP-PEST proteins (e.g represented by this sequence) have shown that  
 CC proline 362 is important for paxillin binding activity. The invention  
 CC relates to a compound that is capable of interfering with the binding of  
 CC PTP-PEST to signalling molecules that are involved in cell migration,  
 CC adhesion or division. The compound can be derived from minimal sequences  
 CC found in binding sites of PTP-PEST. The invention also relates to a  
 CC method for finding a genuine substrate for an enzyme in a cell that  
 CC normally expresses the wild type form of the enzyme. A mutant version of  
 CC PTP-PEST (see AA67251 and AA67253) is used in this method. The  
 CC compounds have cytostatic and antiinflammatory activity. The compounds  
 CC are used for making medicaments for treating a disease related with cell  
 CC proliferation, migration, inflammation and angiogenesis, especially  
 CC cancer. The novel method is used for identifying a genuine substrate for  
 CC an enzyme.  
 XX SQ Sequence 20 AA;  
 Query Match 27.6%; Score 34; DB 21; Length 20;  
 Best Local Similarity 43.8%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 4; Gaps 1;  
 QY 4 HPIPEDSPWTDSTP 19  
 ||:| | :| | |  
 Db 5 hvvp---piltpsap 16  
 RESULT 14  
 AAW24962  
 ID AAW24962 standard; peptide; 14 AA.  
 XX AC AAW24962;  
 DT 28-OCT-1997 (first entry)  
 XX DE C3/C4/C5-derived complement inhibitory peptide C4-B1.  
 XX Indel; insertion; deletion; protein family; sequence alignment;  
 KW interaction site; ligand; interface peptide; binding complex;  
 KW modulation; complement; vaccine; antibody; tissue rejection;  
 KW ischaemia; aneurysm; chronic inflammation; autoimmune disease.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "acylated N-terminus"  
 FT  
 XX PN WO9701578-Al.  
 XX PA 16-JAN-1997.  
 XX PD

XX XX 27-JUN-1996; 96WO-US10958.  
 XX XX 14-JUN-1996; 96US-0663617.  
 PR XX 29-JUN-1995; 95US-0000674.  
 XX XX (MEDI-) MEDICAL BIOLOGY INST.  
 PA Ogata RT;  
 XX WPI; 1997-100166/09.  
 XX Modulatory cpds. identified by screening peptide(s) derived from  
 PT indels - esp. from regions involved in protein-protein interaction,  
 PT useful as inhibitors or activators of the complement system  
 XX PS Claim 30; Page 23; 161pp; English.  
 XX CC The peptides AAW24940-69 represent peptides derived from "indel"  
 CC sequences. Indels are peptide sequences corresponding to insertion  
 CC or deletion sequences in members of a family of proteins when their  
 CC sequences are aligned and compared. The peptides often correspond  
 CC to sites of interaction between proteins or proteins and their ligands.  
 CC The invention relates to a method for identifying regions of contact  
 CC (interface peptides) in protein-protein binding complexes and for the  
 CC preparation of modulators of this interaction. The peptides presented  
 CC here are modulators of the complement system (including peptides that  
 CC act as vaccines by inducing modulating antibodies) which are useful  
 CC therapeutically, e.g to prevent or treat tissue rejection, ischaemia,  
 CC aneurysm, chronic inflammation, autoimmune diseases, etc. This peptide  
 CC is derived from the C3/C4/C5 family of proteins in a region proximal to  
 CC indel-2.  
 XX SQ Sequence 14 AA;  
 Query Match 26.8%; Score 33; DB 18; Length 14;  
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 11 SPWITDS 17  
 ||:| | |  
 Db 8 spwlkds 14  
 RESULT 15  
 AAW25373  
 ID AAW25373 standard; peptide; 16 AA.  
 XX AC AAW25373;  
 XX DT 27-MAR-1998 (first entry)  
 XX DE Src SH3 domain binding peptide SEQ ID NO:157.  
 XX KW Cortactin; SH3 domain; binding peptide; Src homology region 3;  
 KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck;  
 KW ABL; PLCgamma; p53bp2; Crk; Yes; Grb2.  
 XX Synthetic.  
 OS Unidentified.  
 XX WO9730074-Al.  
 XX 21-AUG-1997.  
 XX 14-FEB-1997; 97WO-US02298.  
 XX 16-FEB-1996; 96US-0602999.  
 XX (CYTO-) CYTOGEN CORP.  
 XX (UYNC-) UNIV NORTH CAROLINA.  
 XX

PI Der CJ, Fowlkes DM, Kay BK, Quilliam LA, Rider JE;  
 PI Sparks AB, Thorn JM;  
 XX WPI; 1997-424972/39.  
 DR  
 DR  
 XX Src homology region 3 binding peptide - used to activate Src  
 PT tyrosine kinase(s) and to stimulate immune response by increasing  
 PT production of certain lymphokine(s), e.g. interleukin-1  
 XX  
 XX  
 PS Claim 16; Page 99; 131pp; English.  
 XX  
 CC The present sequence represents a Src homology region 3 (SH3) binding  
 CC peptide. SH3 binding peptides are selected from: (a) peptides which  
 CC bind the SH3 domain of Cortactin; (b) peptides which bind the middle  
 CC SH3 domain of Nck; (c) peptides which bind the SH3 domain of Abl; (d)  
 CC peptides which bind the SH3 domain of Src; (e) peptides which bind the  
 CC SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of  
 CC p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk;  
 CC (h) peptides which bind the SH3 domain of Yes; and (i) peptides which  
 CC bind the amino-terminal SH3 domain of Grb2. The purified binding  
 CC peptides can be used in the method to identify inhibitors of their  
 CC binding to their respective SH3 domains, which could be used to  
 CC modulate the pharmacological activity of proteins or polypeptide  
 CC containing the SH3 domain. The peptides can also be used to activate  
 CC Src or Src-related protein tyrosine kinases, to stimulate the immune  
 CC response by increasing the production of certain lymphokines, e.g.  
 CC tumour necrosis factor-alpha and interleukin-1, or to deliver a  
 CC conjugated molecule to certain cellular compartments containing Src or  
 CC Src related proteins.  
 XX  
 SQ Sequence 16 AA;

Query Match 26.8%; Score 33; DB 18; Length 16;  
 Best Local Similarity 44.4%; Pred. NO. 2.8e+02;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 PIPDESPW 13  
 Db 8 pipetdtpy 16

Search completed: February 21, 2002, 16:42:58  
 Job time: 232 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:43:35 ; Search time 31.57 seconds  
(without alignments)  
14.969 Million cell updates/sec

Title: US-08-753-851-6  
Perfect score: 123  
Sequence: 1 CTVHPIDDESPWITDTPRI 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 114342

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	100.0	21	2	US-08-143-311B-6
2	60	48.8	20	2	US-08-143-311B-7
3	40	32.5	20	4	US-09-230-421-6
4	34	27.6	14	1	US-07-745-382-7
5	34	27.6	14	1	US-07-921-848-7
6	34	27.6	14	1	US-08-165-301A-7
7	34	27.6	14	4	US-08-810-436-7
8	34	27.6	14	5	PCT-US94-14179-7
9	34	27.6	18	2	US-09-017-205-44
10	33	26.8	12	5	PCT-US94-05905-18
11	33	26.8	20	1	US-07-678-974D-31
12	33	26.8	20	2	US-08-934-915-176
13	33	26.8	20	2	US-08-934-915-181
14	33	26.8	20	2	US-08-945-168-36
15	33	26.8	20	4	US-08-602-999A-157
16	33	26.8	21	2	US-08-934-915-174
17	32	26.0	16	4	US-08-602-999A-188
18	32	26.0	19	2	US-08-522-269B-10
19	32	26.0	19	4	US-09-294-923-10
20	32	26.0	21	2	US-08-522-269B-5
21	32	26.0	21	4	US-09-294-923-5
22	31	25.2	12	5	PCT-US94-05905-17
23	31	25.2	21	1	US-08-150-788A-251
24	31	25.2	21	1	US-08-383-474B-254
25	31	25.2	21	1	US-08-465-391A-251
26	31	25.2	21	2	US-08-464-538B-251
27	31	25.2	21	2	US-08-463-076E-308

Sequence 3, Appli  
Sequence 1, Appli  
Sequence 163, App  
Sequence 16, Appl  
Sequence 16, Appl  
Patent No. 5217891  
Sequence 16, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 21, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 32, Appl

28 30.5 24.8 20 2 US-08-617-929-3  
29 30.5 24.8 20 5 PCT-US92-07218-1  
30 30 24.4 16 1 US-07-942-245-163  
31 30 24.4 17 1 US-08-370-567-16  
32 30 24.4 17 1 US-08-438-759-16  
33 30 24.4 17 5 PCT-US94-05684-16  
34 29 23.6 9 6 5217891-8  
35 29 23.6 10 1 US-08-555-394-16  
36 29 23.6 10 3 US-08-745-892-16  
37 29 23.6 14 2 US-08-854-222-11  
38 29 23.6 14 3 US-09-188-039-11  
39 29 23.6 15 1 US-08-277-660A-21  
40 29 23.6 15 1 US-08-424-957-13  
41 29 23.6 15 4 US-09-035-686-13  
42 29 23.6 15 5 PCT-US93-06751-32  
43 29 23.6 17 1 US-08-451-947-32  
44 29 23.6 17 2 US-08-424-826A-32  
45 29 23.6 17 3 US-08-928-694-32

## ALIGNMENTS

RESULT 1  
US-08-143-311B-6  
; Sequence 6, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-6

Query Match 100.0%; Score 123; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. NO. 2.8e-11;  
Matches 21; Conservative 0; Mismatches 0; Indels

QY 1 CTVHPIPEDSPWITDSTPRI 21  
 |||||  
 Db 1 CTVHPIPEDSPWITDSTPRI 21

## RESULT 2

US-08-143-311B-7  
; Sequence 7, Application US/08143311B  
: Patent No. 5863540

```

:
: GENERAL INFORMATION:
: APPLICANT: HAINES, BARTON F.
: APPLICANT: HALE, LAURA P.
: APPLICANT: PATTON, KAREN L.
: APPLICANT: TELEN, MARILYN J.
: APPLICANT: LIAO, HUA-XIN
: TITLE OF INVENTION: AN ADHESION MOLECULE
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
:

```

NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714

```

,
,
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.25
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/143,311B
, FILING DATE: 29-OCT-1993
,

```

CLASSIFICATION: 436  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/973,339  
 FILING DATE: 30-OCT-1992  
 CLASSIFICATION: 436  
 APPLICATION NUMBER: 07/669,730  
 FILING DATE: 15-MAR-1991  
 CLASSIFICATION: 436  
 ATTORNEY/AGENT INFORMATION:

NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 7

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;

```

Query Match 48.8%; Score 60; DB 2; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 1; Indels

Qy	10	DSPWITDSTPRI	21
Db	1	DSPWITDSTPRI	12

### RESULT 3

```

US-09-230-421-6
; Sequence 6, Application US/09230421
; Patent No. 6200577
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-HERPESVIRAL AGENTS AND ASSAYS
; TITLE OF INVENTION: THEREFOR

```

```

; FILE OF INVENTION: 18189C
; FILE REFERENCE: P18189C
; CURRENT APPLICATION NUMBER: US/09/230,421
; CURRENT FILING DATE: 1999-01-25

```

```

; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 20

```

```
;
;
; ORGANISM: Artificial Sequence
; FEATURE:
```

7. FEATURE: ; OTHER INFORMATION: SYNTHETIC PEPTIDES DERIVED FROM THE VP22TRUNC  
; OTHER INFORMATION: SEQUENCE  
US-09-230-421-6

Query Match 32.5%; Score 40; DB 4; Length 20;  
Best Local Similarity 46.7%; Pred. No. 11;  
Matches 7; Conservative 3; Mismatches 1; Indels

```

QY      7 PDEDSPWITDSTPRI 21
        1:|:|:|  | | |
Db      6 PNPDA PW---TPRV 16

```

## RESULTS

US-07-745-382-7  
; Sequence 7, Application US/07745382  
; Patent No. 5270181

/ PATENT NO.: 5270101  
 / GENERAL INFORMATION:  
 / APPLICANT: McCoy, John  
 / APPLICANT: McVallie, Edward  
 / TITLE OF INVENTION: Peptide and Protein Fusions To  
 / TITLE OF INVENTION: Thioresoxin and Thioresoxin-Like Molecules  
 / NUMBER OF SEQUENCES: 24  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Genetics Institute, Inc.  
 / STREET: 87 CambridgePark Drive  
 / CITY: Cambridge  
 / STATE: Massachusetts  
 / COUNTRY: U.S.A.

```
?
? ZIB# 02140
? COMPUTER TYPE: FLOPPY disk
? MEDIUM TYPE: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/745,382
? FILING DATE: 19910814
?
```

/ CLASSIFICATION: 435  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: US 07/652,531  
 / FILING DATE: 06-FEB-1991  
 / ATTORNEY/AGENT INFORMATION:

NAME: Cseri, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: G15188A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 876-1170  
TELEFAX: (617) 876-5851

```

; LEEBAX. (017) 870 5051
;
; INFORMATION FOR SEQ ID NO: 7:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 14 amino acids
;
; TYPE: AMINO ACID
;
; TOPOLOGY: unknown
;
; MOLECULE TYPE: peptide
;

```

US-07-745-382-7

Query Match 27.6%; Score 34; DB 1; Length 14;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 12 PWITDSTP 19  
||| :||  
Db 1 PWINGATP 8

RESULT 5

US-07-921-848-7  
; Sequence 7, Application US/07921848  
; Patent No. 5292646  
; GENERAL INFORMATION:  
; APPLICANT: McCoy, John  
; APPLICANT: DiBlasio-Smith, Edward  
; APPLICANT: Grant, Kathleen  
; TITLE OF INVENTION: Peptide and Protein Fusions To  
; TITLE OF INVENTION: Thioresoxin and Thioresoxin-Like Molecules  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/921,848  
; FILING DATE: 19920728  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/652,531  
; FILING DATE: 06-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/745,382  
; FILING DATE: 14-AUG-1991

ATTORNEY/AGENT INFORMATION:  
; NAME: Cseri, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: G15188A  
; TELEPHONE: (617) 876-1170  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide

US-07-921-848-7

Query Match 27.6%; Score 34; DB 1; Length 14;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 12 PWITDSTP 19  
||| :||  
Db 1 PWINGATP 8

RESULT 6

US-08-165-301A-7  
; Sequence 7, Application US/08165301A  
; Patent No. 5646016

GENERAL INFORMATION:  
; APPLICANT: McCoy, John  
; APPLICANT: DiBlasio-Smith, Elizabeth  
; APPLICANT: Grant, Kathleen  
; APPLICANT: LaVallie, Edward R.  
; TITLE OF INVENTION: PEPTIDE AND PROTEIN FUSIONS TO  
; TITLE OF INVENTION: THIOREDOXIN, THIOREDOXIN-LIKE MOLECULES, AND MODIFIED  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/165,301A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meinert, M. C.  
; REGISTRATION NUMBER: 33,544  
; REFERENCE/DOCKET NUMBER: G1 5188D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 876-1170  
; TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-165-301A-7

Query Match 27.6%; Score 34; DB 1; Length 14;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 12 PWITDSTP 19  
||| :||  
Db 1 PWINGATP 8

RESULT 7

US-08-810-436-7  
; Sequence 7, Application US/08810436  
; Patent No. 6143524  
; GENERAL INFORMATION:

APPLICANT: McCoy, John  
; APPLICANT: DiBlasio-Smith, Elizabeth  
; APPLICANT: Grant, Kathleen  
; APPLICANT: LaVallie, Edward R.  
; TITLE OF INVENTION: PEPTIDE AND PROTEIN FUSIONS TO  
; TITLE OF INVENTION: THIOREDOXIN, THIOREDOXIN-LIKE MOLECULES, AND MODIFIED  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk



; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/810,436  
; FILING DATE: 04-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/165,301  
; FILING DATE: 10-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meinert, M. C.  
; REGISTRATION NUMBER: 33,544  
; REFERENCE/DOCKET NUMBER: GI 5188D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 876-1170  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-810-436-7

Query Match 27.6%; Score 34; DB 4; Length 14;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 PWITDSTP 19  
DB 1 PWINGATP 8

RESULT 8  
PCT-US94-14179-7  
; GENERAL INFORMATION:  
; APPLICANT: McCoy, John  
; APPLICANT: DiBlasio-Smith, Elizabeth  
; APPLICANT: Grant, Kathleen  
; APPLICANT: Lavallie, Edward R.  
; TITLE OF INVENTION: PEPTIDE AND PROTEIN FUSIONS TO  
; TITLE OF INVENTION: THIOREDOXIN, THIOREDOXIN-LIKE MOLECULES, AND MODIFIED  
; TITLE OF INVENTION: THIOREDOXIN-LIKE MOLECULES  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/14179  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meinert, M. C.  
; REGISTRATION NUMBER: 33,544  
; REFERENCE/DOCKET NUMBER: GI 5188D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 876-1170  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; PCT-US94-14179-7

Query Match 27.6%; Score 34; DB 5; Length 14;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 PWITDSTP 19  
DB 1 PWINGATP 8

RESULT 9  
US-09-017-205-44  
; Sequence 44, Application US/09017205  
; Patent No. 5965357  
; GENERAL INFORMATION:  
; APPLICANT: Marsden, Howard S  
; TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN  
; TITLE OF INVENTION: DIAGNOSIS OF HERPES SIMPLEX VIRUS TYPE 2  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye PC  
; STREET: 8th Floor, 1100 No. 5965357th Glebe Road  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/017,205  
; FILING DATE: 02-FEB-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mitchard, Leonard C  
; REGISTRATION NUMBER: 29,009  
; REFERENCE/DOCKET NUMBER: 604-436  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)816-4000  
; TELEFAX: (703)816-4100  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide from HSV-2 glycoprotein G  
; FRAGMENT TYPE: internal  
; US-09-017-205-44

Query Match 27.6%; Score 34; DB 2; Length 18;  
Best Local Similarity 43.8%; Pred. No. 70;  
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 PIPEDSPWITDSTPR 20  
DB 1 PAPTTPPTTSTHATPR 16

RESULT 10  
PCT-US94-05905-18  
; Sequence 18, Application PC/TUS9405905  
; GENERAL INFORMATION:  
; APPLICANT:

```

; TITLE OF INVENTION:  TRNA BINDING-DEPENDENT INHIBITION OF MICROBIAL
; TITLE OF INVENTION:  PATHOGEN GROWTH
; NUMBER OF SEQUENCES:  22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
; STREET:  Two Militia Drive
; CITY:  Lexington
; STATE:  MA
; COUNTRY:  USA
; ZIP:  02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  PCT/US94/05905
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  US 08/068,382
; FILING DATE:  28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME:  Brook, David E.
; REGISTRATION NUMBER:  22,592
; REFERENCE/DOCKET NUMBER:  MIT-6299A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  617-861-6240
; TELEFAX:  617-861-9540
; TELEX:  951794
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US94-05905-18

Query Match      26.8%; Score 33; DB 5; Length 12;
Best Local Similarity 57.1%; Pred. No. 61;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      8 DEDSPWI 14
DB      5 DEQAPWV 11

RESULT 11
US-07-678-974D-31
; Sequence 31, Application US/07678974D
; Patent No. 5629146
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; TITLE OF INVENTION:  METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
; NUMBER OF SEQUENCES:  67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  BERMAN & AISENBERG
; STREET:  1730 RHODE ISLAND AVENUE, N.W.,
; CITY:  WASHINGTON
; STATE:  D.C.
; COUNTRY:  U.S.A.
; ZIP:  20036-3186
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/07/678,974D
; FILING DATE:  25-JUN-1991
; CLASSIFICATION:  435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME:  AISENBERG, Irwin M.
; REGISTRATION NUMBER:  19,007
; REFERENCE/DOCKET NUMBER:  SG19171
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  202-293-1404
; TELEFAX:  202-872-0493
; TELEX:  440 069 AIS UI
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-07-678-974D-31

Query Match      26.8%; Score 33; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY      1 CTVH--PIPEDSP 12
DB      2 CQKHTPPAPKEDDP 15

RESULT 12
US-08-934-915-176
; Sequence 176, Application US/08934915
; Patent No. 5932412
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; APPLICANT: CHENG, HWEE-MING
; TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
; TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
; TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
; TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
; TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
; NUMBER OF SEQUENCES:  193
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  MASON & ASSOCIATES, P.A.
; STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
; CITY:  CLEARWATER
; STATE:  FLORIDA
; COUNTRY:  U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  Windows 3.0
; SOFTWARE:  Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/934,915
; FILING DATE:  22-SEP-1997
; CLASSIFICATION:  435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  07/949,836
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:  LOUISE A. Foutch
; REGISTRATION NUMBER:  37,133
; REFERENCE/DOCKET NUMBER:  1946.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  813-538-3800
; TELEFAX:  813-538-3820
; TELEX:
; INFORMATION FOR SEQ ID NO: 176:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:  peptide
; US-08-934-915-176

```

Query Match 26.8%; Score 33; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 1;

QY 1 CTVH--PIPEDSP 12  
| | | | |  
Db 2 CQKHTPPAPKDDP 15

## RESULT 13

US-08-934-915-181  
; Sequence 181, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, JENA  
; APPLICANT: CHENG, HWEI-MING  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. Foutch  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
TELEX:

INFORMATION FOR SEQ ID NO: 181:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-181

Query Match 26.8%; Score 33; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 1;

QY 1 CTVH--PIPEDSP 12  
| | | | |  
Db 2 CQKHTPPAPKDDP 15

## RESULT 14

US-08-945-168-36  
; Sequence 36, Application US/08945168  
; Patent No. 5989548  
; GENERAL INFORMATION:

APPLICANT: DILLNER, JOAKIM  
TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA  
TITLE OF INVENTION: VIRUS  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945,168  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE96/00533  
FILING DATE: 23-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9501512-9  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-945-168-36

Query Match 26.8%; Score 33; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 1;

QY 1 CTVH--PIPEDSP 12  
| | | | |  
Db 2 CQKHTPPAPKDDP 15

## RESULT 15

US-08-602-999A-157  
; Sequence 157, Application US/08602999A  
; Patent No. 6184205  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York

COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999A  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 157:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-602-999A-157

Query Match 26.8%; Score 33; DB 4; Length 20;  
Best Local Similarity 44.4%; Pred. No. 1.1e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 PIPDEDSPW 13  
I: I: I: I:  
Db 10 PLPETDTPY 18

Search completed: February 21, 2002, 16:43:36  
Job time: 170 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:44:16 ; Search time 33.6 Seconds  
(without alignments)  
47.609 Million cell updates/sec

Title: US-08-753-851-7

Perfect score: 113

Sequence: 1 DSPWITDTRIFATRDQDTI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4046

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	31.9	20	S00492	hemocyanin chain I
2	30	26.5	18	A35678	hypothetical prote
3	29	25.7	16	S02473	coat protein VP1 -
4	29	25.7	16	A48301	glutamate-1-semial
5	27	23.9	16	S66613	protein p12E - Fri
6	27	23.9	19	S60633	H+-transporting AT
7	27	23.9	21	S78416	ribosomal protein
8	26	23.0	20	S10876	hypothetical prote
9	26	23.0	21	F28027	protein P9 - curle
10	25	22.1	21	S08590	NADH dehydrogenase
11	24	21.2	15	PX0031	mixed lymphocyte r
12	24	21.2	19	PQ0548	capsid protein VP2
13	24	21.2	20	S77981	cytochrome-c oxida
14	24	21.2	20	A44773	pollen allergen I
15	23	20.4	17	C85956	hypothetical prote
16	23	20.4	20	I49423	cytotoxic T-lympho
17	23	20.4	20	A31049	calsequestrin, fas
18	22	19.5	7	PJ0029	pev-kinin 1 - pena
19	22	19.5	9	JD0302	xenopsin-related p
20	22	19.5	9	A60320	xenopsin-related p
21	22	19.5	11	P71304	amine oxidase (cop
22	22	19.5	11	P00334	dextranucrase (EC
23	22	19.5	13	I51905	collecting duct wa
24	22	19.5	16	A20190	hypodermin B - ear
25	22	19.5	16	B23692	transcription fact
26	22	19.5	17	1 GMSH	gastrin - sheep
27	22	19.5	17	D48138	d(TTAGGG)n-binding
28	22	19.5	19	PQ0790	NADH dehydrogenase
29	22	19.5	20	A56899	serum heterodimer,

30 22 19.5 21 2 A38837 T-cell receptor be  
31 22 19.5 21 2 D61497 seed protein ws-19  
32 21 18.6 10 2 PH0895 T-cell receptor be  
33 21 18.6 12 2 PT0274 Ig heavy chain CRD  
34 21 18.6 12 2 S21163 NAD+ ADP-ribosyltr  
35 21 18.6 13 2 PC2369 unidentified 85K p  
36 21 18.6 13 2 S47356 T-cell antigen rec  
37 21 18.6 13 2 S47385 T-cell antigen rec  
38 21 18.6 15 2 S42741 ubiquinol--cytochr  
39 21 18.6 16 2 PH1473 T-cell receptor be  
40 21 18.6 18 2 C49254 tCr C gamma 1 chai  
41 21 18.6 20 2 A56045 urinary tract ston  
42 21 18.6 20 2 I46652 T-cell receptor de  
43 21 18.6 21 2 PQ0119 dystrophin - pig (  
44 21 18.6 21 2 A45761 olfactomedin - bul  
45 20.5 18.1 18 2 G02018 proteasome chain L

#### ALIGNMENTS

RESULT 1

S00492 hemocyanin chain Ia - Japanese spiny lobster (fragment)  
C:Species: Panulirus japonicus (Japanese spiny lobster)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 01-Aug-1997  
C:Accession: S00492  
R:Makino, N.; Kimura, S.  
Eur. J. Biochem. 173, 423-430, 1988  
A:Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.  
A:Reference number: S00492; MUID:88196131  
A:Accession: S00492  
A:Molecule type: protein  
A:Residues: 1-20 <MAK>  
C:Superfamily: hemocyanin  
C:Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 31.9%; Score 36; DB 2; Length 20;  
Best Local Similarity 46.2%; Pred. No. 18;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 DSTDRIFATRDQD 19

||||:|:|

Db 7 DSTDKLLAQKQDD 19

RESULT 2

A35678 hypothetical protein (proenkephalin 5' region) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 05-Nov-1999  
C:Accession: A35678  
R:Kilpatrick, D.L.; Zinn, S.A.; Fitzgerald, M.; Higuchi, H.; Sabol, S.L.; Meyerhardt, M.; Cell. Biol. 10, 3717-3726, 1990  
A:Title: Transcription of the rat and mouse proenkephalin genes is initiated at distal  
A:Reference number: A35678; MUID:90287163  
A:Accession: A35678  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-18 <KIL>  
A:Cross-references: GB:M55181; NID:g201032; PIDN:AAA40127.1; PID:g201033

Query Match 26.5%; Score 30; DB 2; Length 18;  
Best Local Similarity 58.3%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

Qy 1 DSPWITDTRDI 12

|||||

Db 7 DSPW----EDRI 14

```

RESULT 3
S02473
coat protein VP1 - human poliovirus 1 (fragment)
C:Species: human poliovirus 1
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 18-Jun-1993
C:Accession: S02473
R:Martin, A.; Wychocki, C.; Couderc, T.; Crainic, R.; Hogle, J.; Girard, M.
EMBO J. 7, 2839-2847, 1988
A:Title: Engineering a poliovirus type 2 antigenic site on a type 1 capsid results in a
A:Reference number: S02473; MUID:89030650
A:Accession: S02473
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-16 <MAR>

Query Match 25.7%; Score 29; DB 2; Length 16;
Best Local Similarity 38.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DSPWITDSTDRIF 13
I : I : I : I
D 4 DNPASTNKDKLF 16

RESULT 4
A48301
glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) - Synechococcus sp. (PCC 6301) (fr
C:Species: Synechococcus sp.
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C:Accession: A48301
R:Grimm, B.; Bull, A.; Welinder, K.G.; Gough, S.P.; Kannangara, C.G.
Carlsberg Res. Commun. 54, 67-79, 1989
A:Title: Purification and partial amino acid sequence of the glutamate 1-semialdehyde am
A:Reference number: A48301; MUID:89374545
A:Accession: A48301
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-16 <GRI>
C:Keywords: intramolecular transferase; isomerase

Query Match 25.7%; Score 29; DB 2; Length 16;
Best Local Similarity 46.2%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 SPWITDSTDRIFA 14
I : I : I : I
D 4 NPFKTIKSDIEIFA 16

RESULT 5
S66613
protein p12E - Friend murine leukemia virus (fragments)
C:Species: Friend murine leukemia virus
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C:Accession: S66613
R:Hensel, J.; Hintz, M.; Karas, M.; Linder, D.; Stahl, B.; Geyer, R.
Eur. J. Biochem. 232, 373-380, 1995
A:Title: Localization of the palmitoylation site in the transmembrane protein p12E of F
A:Reference number: S66613; MUID:96035869
A:Accession: S66613
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-9;10-16 <HEN>

Query Match 23.9%; Score 27; DB 2; Length 16;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SPWIT 6
I : I : I

```

```

Db 10 SPWFT 14

RESULT 6
S60633
H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - brine shrimp mitochondrion (st
C:Species: mitochondrion Artemia sp. (brine shrimp)
A:Variety: strain La Mata
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-Dec-1999
C:Accession: S60633
R:Perez, M.L.; Valverde, J.R.; Batuecas, B.; Amat, F.; Marco, R.; Garesse, R.
J. Mol. Evol. 38, 156-168, 1994
A:Title: Speciation in the Artemia genus: mitochondrial DNA analysis of bisexual and
A:Reference number: S60624; MUID:94223692
A:Accession: S60633
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-19 <PER>
A:Cross-references: EMBL:X67263; NID:g11210; PIDN:CAAA47685.1; PID:g11211
A:Experimental source: strain La Mata
A:Note: the source is designated as Artemia parthenogenetica
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
C:Genetics:
A:Gene: ATP8
A:Genome: mitochondrion
A:Genetic code: SGC4
C:Superfamily: H+-transporting ATP synthase protein 8
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion;

Query Match 23.9%; Score 27; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PWIT 6
I : I : I : I
D 8 PWIT 11

RESULT 7
S78416
ribosomal protein RL28, mitochondrial [validated] - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000
C:Accession: S78416
R:Goldschmidt-Reisin, S.; Graack, H.R.
submitted to the Protein Sequence Database, February 1998
A:Reference number: S78411
A:Accession: S78416
A:Molecule type: protein
A:Residues: 1-21 <GOL>
A:Note: the protein is designated as mitochondrial ribosomal protein L28
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 23.9%; Score 27; DB 2; Length 21;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 DRIFATRDQD 19
I : I : I : I
D 9 DKVFQRPED 18

RESULT 8
S10876
hypothetical protein - human
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Nov-1994
C:Accession: S10876
R:Assouline, Z.; Kerbirou-Nabias, D.M.; Pietu, G.; Thomas, N.; Bahnak, B.R.; Meyer,
Biochem. Biophys. Res. Commun. 153, 1159-1166, 1988
A:Title: The human gene for von Willebrand factor. Identification of repetitive Alu s

```

A:Reference number: S10876; MUID:88268889  
 A:Accession: S10876  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-20 <ASS>  
 A:Cross-references: EMBL:X07258

Query Match 23.0%; Score 26; DB 2; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 6.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SPWIT 6  
 |||  
 Db 16 SPWAT 20

RESULT 9  
 F28027

protein p9 - curled-leaved tobacco (fragment)  
 C:Species: Nicotiana glauca (curled-leaved tobacco)  
 C:Date: 19-May-1989 #sequence\_revision 19-May-1989 #text\_change 18-Jun-1993  
 C:Accession: F28027  
 R:Bauw, G.; De Loose, M.; Inze, D.; Van Montagu, M.; Vandekerckhove, J.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 4806-4810, 1987  
 A:Title: Alterations in the phenotype of plant cells studied by NH2-terminal amino acid-  
 A:Reference number: A94167  
 A:Accession: F28027  
 A:Molecule type: protein  
 A:Residues: 1-21 <BAU>

Query Match 23.0%; Score 26; DB 2; Length 21;  
 Best Local Similarity 38.5%; Pred. No. 7.3e+02;  
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SPWITSDRIFA 14  
 |||  
 Db 6 SPVLDITNGDKIYA 18

RESULT 10  
 S08590

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Debney's tobacco chloroplast (fra  
 C:Species: Chloroplast Nicotiana glauca (Debney's tobacco)  
 C:Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 11-Jun-1999  
 C:Accession: S08590  
 R:Hayashida, N.; Matsubayashi, T.; Shinozaki, K.; Suglura, M.; Inoue, K.; Hiyama, T.  
 Curr. Genet. 12, 247-250, 1987  
 A:Title: The gene for the 9 kd polypeptide, a possible apoprotein for the iron-sulfur ce  
 A:Reference number: S07170; MUID:88210537  
 A:Accession: S08590  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-21 <HAY>

A:Cross-references: EMBL:X05881; NID:g11790; PIDN:CAA29305.1; PID:g11793  
 C:Genetics:

A:Genome: chloroplast  
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
 C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 22.1%; Score 25; DB 2; Length 21;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PWIT 6  
 |||  
 Db 14 PWLT 17

RESULT 11

PX0031  
 mixed lymphocyte reaction inhibitor - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 24-Feb-1995  
 C:Accession: PX0031  
 R:Shinomiyama, T.; Ohara, T.; Wada, N.; Omori, A.; Kamada, N.  
 J. Biochem. 107, 435-439, 1990  
 A:Title: Rat liver arginase suppresses mixed lymphocyte reaction.  
 A:Reference number: PX0031; MUID:90256720  
 A:Accession: PX0031  
 A:Molecule type: protein  
 A:Residues: 1-15 <SHI>  
 A:Experimental source: liver  
 C:Keywords: lymphocyte

Query Match 21.2%; Score 24; DB 2; Length 15;  
 Best Local Similarity 33.3%; Pred. No. 1e+03;  
 Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DSPWITDST 9  
 : ||| :  
 Db 1 EEPWMSMS 9

RESULT 12

PQ0548  
 capsid protein VP26 - human herpesvirus 1 (fragment)  
 C:Species: human herpesvirus 1  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
 C:Accession: PQ0548  
 R:Davidson, M.D.; Rixon, F.J.; Davison, A.J.  
 J. Gen. Virol. 73, 2709-2713, 1992  
 A:Title: Identification of genes encoding two capsid proteins (VP24 and VP26) of herp  
 A:Reference number: PQ0544; MUID:93019027  
 A:Accession: PQ0548  
 A:Molecule type: protein  
 A:Residues: 1-19 <DAV>  
 A:Experimental source: strain 17  
 C:Genetics:  
 A:Gene: UL35  
 C:Keywords: capsid protein

Query Match 21.2%; Score 24; DB 2; Length 19;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 12 IFATRDQD 19  
 : || :  
 Db 4 VLATNNQD 11

RESULT 13

S77981  
 cytochrome-c oxidase (EC 1.9.3.1) chain Va.1 - bigeye tuna (fragment)  
 C:Species: Thunnus obesus (bigeye tuna)  
 C:Date: 17-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Feb-1998  
 C:Accession: S77981  
 R:Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.  
 submitted to the Protein Sequence Database, June 1997  
 A:Reference number: S77980  
 A:Accession: S77981

A:Molecule type: protein  
 A:Residues: 1-20 <ARN>  
 A:Experimental source: heart; liver  
 C:Genetics:  
 A:Genome: nuclear

C:Function:  
 A:Pathway: oxidative phosphorylation; respiratory chain  
 C:Superfamily: mammalian cytochrome-c oxidase chain Va  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 21.2%; Score 24; DB 2; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSPWIT 6  
 | : | : |  
 DB 12 DARWVT 17

## RESULT 14

A44773  
 pollen allergen I - Japanese cedar (fragment)  
 C:Species: Cryptomeria japonica (Japanese cedar)  
 C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 11-Jan-2000  
 C:Accession: A44773  
 R:Tanihara, M.; Ando, S.; Usui, M.; Kurimoto, M.; Sakaguchi, M.; Inouye, S.; Matuhasi, T.  
 FEBS Lett. 239, 329-332, 1988  
 A:Title: N-terminal amino acid sequence of a major allergen of Japanese cedar pollen (Cr  
 A:Reference number: A44773; MUID:89031257  
 A:Accession: A44773  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <TAN>  
 C:Superfamily: pectate lyase LAT59  
 C:Keywords: pollen

Query Match 21.2%; Score 24; DB 2; Length 20;  
 Best Local Similarity 62.5%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSPWITDS 8  
 | | | | |  
 DB 5 DSXWRGDS 12

## RESULT 15

C85956  
 hypothetical protein Z4327 [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: C85956  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: AB5480; MUID:21074935; PMID:11206551  
 A:Accession: C85956  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-17 <STO>  
 A:Cross-references: GB:AE005174; NID:gi2517535; PIDN:AA058111.1; GSPDB:GN00145; UWGP:Z43  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z4327

Query Match 20.4%; Score 23; DB 2; Length 17;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 STDRIAT 15  
 | | | | |  
 DB 6 STDDEFIT 13

Search completed: February 21, 2002, 16:44:16  
 Job time: 190 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:52:21 ; Search time 20.15 Seconds  
(without alignments)  
38.211 Million cell updates/sec

Title: US-08-753-851-7  
Perfect score: 113  
Sequence: 1 DSPWITDSTRIFATRDQDTI 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1160

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	29	25.7	19 1	FTBB_VULVU
2	24	21.2	20 1	COXA_THUOB
3	23	20.4	20 1	CAQS_RAT
4	21.5	19.0	19 1	H170_RAT
5	21	18.6	15 1	PGKH_PHYPA
6	21	18.6	16 1	LE06_BIOGL
7	21	18.6	16 1	PPAC_BACME
8	20.5	18.1	20 1	APAL_ERYP
9	20	17.7	17 1	GAST_MACMU
10	20	17.7	18 1	FTBB_ANAPL
11	19	16.8	9 1	UHA2_HUMAN
12	19	16.8	10 1	TOPL_SALTY
13	19	16.8	13 1	UHA3_CANFA
14	19	16.8	14 1	MY14_PHEVI
15	19	16.8	14 1	TKNM_RANNA
16	19	16.8	15 1	OBPA_MAMBR
17	19	16.8	17 1	ATPL_PAVLU
18	19	16.8	18 1	RL23_HALCU
19	18	15.9	11 1	ANGT_CRIGE
20	18	15.9	12 1	NO40_SOYBN
21	18	15.9	12 1	ULAL_MOUSE
22	18	15.9	13 1	LPAA_PORGI
23	18	15.9	14 1	CAT2_FASHE
24	18	15.9	15 1	MR2A_PALPR
25	18	15.9	16 1	BAIT_EUBSP
26	18	15.9	16 1	MR2B_PALPR
27	18	15.9	16 1	NK3_PALPR
28	18	15.9	18 1	HEMH_THETS
29	18	15.9	19 1	FTBB_TAPTE
30	18	15.9	20 1	CATA_ACIRA
31	18	15.9	20 1	ELAS_GADMO
32	18	15.9	20 1	LEC2_MACPO
33	18	15.9	20 1	LEC3_ARTIN

34 18 15.9 20 1 LEC3\_MACPO  
35 18 15.9 21 1 BTX\_ATRBI  
36 18 15.9 21 1 LEC1\_MACPO  
37 18 15.9 21 1 LEC2\_ARTIN  
38 18 15.9 21 1 SRD2\_ATREN  
39 18 15.9 21 1 TRYP\_PROAT  
40 18 15.9 21 1 YFLA\_METVO  
41 17 15.0 8 1 ANG2\_BOTJA  
42 17 15.0 10 1 ANG1\_BOTJA  
43 17 15.0 10 1 ANG2\_BOVIN  
44 17 15.0 10 1 ANG1\_CHICK  
45 17 15.0 11 1 NUHM\_CANFA

## ALIGNMENTS

RESULT 1  
FTBB\_VULVU  
ID FTBB\_VULVU STANDARD; PRT; 19 AA.  
AC P14482;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-JAN-1990 (Rel. 13, Last annotation update)  
DE FIBRINOPEPTIDE B.  
OS Vulpes vulpes (Red fox).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Vulpes.  
OX NCBI\_TaxID=9627;  
RN [1]  
RP SEQUENCE.  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COPOLYMER IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
CC INTERPRO: IPR002181; Fibrinogen\_C.  
DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma.  
FT NON\_TER 19  
SQ SEQUENCE 19 AA; 2305 MW; 0B0D756C24A1F2C0 CRC64;

Query Match 25.7%; Score 29; DB 1; Length 19;  
Best Local Similarity 50.0%; Pred. No. 86;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 TDSTDRIFATRD 17  
|| :|| :|  
Db 6 TDEERIVSTVD 17

RESULT 2  
COXA\_THUOB  
ID COXA\_THUOB STANDARD; PRT; 20 AA.  
AC P80372;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-1 (EC 1.9.3.1) (FRAGMENT).  
OS Thunnus obesus (Bigeye tuna).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;  
OC Scombridae; Thunnus.  
OX NCBI\_TaxID=8241;

```

RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RA MEDLINE=97454291; PubMed=9310366;
RX Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2404 MW; 7E82E43B7157355E CRC64;

Query Match 21.2%; Score 24; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 6e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DSPWIT 6
Db 12 DARWVT 17

RESULT 3
CAQS_RAT STANDARD; PRT; 20 AA.
ID CAQS_RAT
AC P19633;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE CALSEQUESTRIN, SKELETAL MUSCLE ISOFORM (ASPARTACTIN) (LAMININ-BINDING
DE PROTEIN) (FRAGMENT).
GN CASQ1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=88331073; PubMed=3417768;
RA Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;
RT "Isolation and characterization of a laminin-binding protein from rat
RT and chick muscle.";
RL J. Cell Biol. 107:687-697(1988).
CC -!- FUNCTION: CALSEQUESTRIN IS A HIGH-CAPACITY, MODERATE AFFINITY,
CC CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE
CC IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH
CC A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40
CC TO 50 MOLES OF CALCIUM. ALSO BINDS LAMININ.
CC -!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE
CC SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF
CC FAST SKELETAL MUSCLE CELLS. ASPARTACTIN IS FOUND IN THE BASAL
CC LAMINA SURROUNDING INDIVIDUAL MUSCLE FIBERS.
CC -!- TISSUE SPECIFICITY: SKELETAL AND HEART MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
DR PIR: A31049; A31049.
DR InterPro: IPR001393; Calsequestrin. PARTIAL.
DR PROSITE: PS00864; CALSEQUESTRIN_2. PARTIAL.
DR PROSITE: PS00863; CALSEQUESTRIN_1.
KW Muscle; Glycoprotein; Calcium-binding.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2238 MW; 92ADE04FC2A69280 CRC64;

Query Match 20.4%; Score 23; DB 1; Length 20;
Best Local Similarity 36.4%; Pred. No. 8.7e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 7 DSTDRIFATRD 17
Db 10 DGVDRVINAD 20

RESULT 4
HI70_RAT STANDARD; PRT; 19 AA.
ID HI70_RAT
AC P21794;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE HORMONE-INDUCED PROTEIN 70 KDA (HIP-70) (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=90208308; PubMed=2181662;
RA Mobbs C.V., Fink G., Pfaff D.W.;
RT "HIP-70: a protein induced by estrogen in the brain and LH-RH in the
RT pituitary.";
RL Science 247:1477-1479(1990).
CC -!- INDUCTION: MOST PROMINENT PROTEIN INDUCED BY ESTROGEN IN
CC HYPOTHALAMUS AND MOST PROMINENT PROTEIN INDUCED BY LH-RH IN
CC PITUITARY.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2113 MW; F793A98720B68E3C CRC64;

Query Match 19.0%; Score 21.5; DB 1; Length 19;
Best Local Similarity 35.3%; Pred. No. 1.4e+03;
Matches 6; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 5 ITDSTDIFATR-DQDT 20
Db 2 VLELTDFENFVSST 18

RESULT 5
PGKH_PHYPA STANDARD; PRT; 15 AA.
ID PGKH_PHYPA
AC P80659;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOGLYCERATE KINASE, CHLOROPLAST (EC 2.7.2.3) (FRAGMENT).
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=32118;
RN [1]
RP SEQUENCE.
RC TISSUE=Protonema; PubMed=9129336;
RX MEDLINE=97275459; PubMed=9129336;
RA Kasten B., Buck F., Nuske J., Reski R.;
RT "Cytokinin affects nuclear- and plastome-encoded energy-converting
RT plastid enzymes.";
RL Planta 201:261-272(1997).
CC -!- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE = ADP +
CC 3-PHOSPHO-D-GLYCEROYL PHOSPHATE.
CC -!- PATHWAY: CALVIN CYCLE.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
DR InterPro: IPR001576; PGK.
DR PROSITE: PS00111; GLYCERATE_KINASE; PARTIAL.
KW Transferase; Kinase; Multigene family; Calvin cycle; Chloroplast.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1531 MW; 56A5EC1F677EBC6 CRC64;

```

Query Match 18.6%; Score 21; DB 1; Length 15;  
 Best Local Similarity 23.1%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 IYDSTDRIATRD 17  
 :|: :|: |  
 Db 2 LTEQASKVALTAD 14

RESULT 6  
 LE06\_BIOGL  
 ID LE06\_BIOGL STANDARD; PRT; 16 AA.  
 AC P80745;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HEMOLYMPH 65 KDA LECTIN BG06 (FRAGMENT).  
 GN BG06.  
 OS Biomphalaria glabrata (Bloodfluke planorb).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Planorbidae; Biomphalaria.  
 OX NCBI\_TaxID=6526;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=M-LINE; TISSUE=Hemolymph;  
 RX MEDLINE=97385165; PubMed=9238039;  
 RA Adema C.M., Hertel L.A., Miller R.D., Loker E.S.;  
 RT "A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:8691-8696(1997).  
 CC -!- FUNCTION: BINDS AND PRECIPITATES ANTIGENS OF THE PARASITE ECHINOSTOMA PARANSEI.  
 CC -!- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.  
 CC -!- INDUCTION: BY INFECTION.  
 KW Lectin.  
 FT NON\_TER 1 16  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1964 MW; A1665754589EF82C CRC64;

Query Match 18.6%; Score 21; DB 1; Length 16;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 FATRDQD 19  
 | | | |  
 Db 2 FTTFDKD 8

RESULT 7  
 PPAC\_BACME  
 ID PPAC\_BACME STANDARD; PRT; 16 AA.  
 AC P56948;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MANGANESE-DEPENDENT INORGANIC PYROPHOSPHATASE (EC 3.6.1.1)  
 DE (PYROPHOSPHATE PHOSPHO-HYDROLASE) (PPASE) (FRAGMENT).  
 GN PPAC.  
 OS Bacillus megaterium.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1404;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=98455825; PubMed=9782505;  
 RA Young T.W., Kuhn N.J., Wadeson A., Ward S., Burges D., Cooke G.D.;  
 RT "Bacillus subtilis ORF yybQ encodes a manganese-dependent inorganic pyrophosphatase with distinctive properties: the first of a new class of soluble pyrophosphatase?";  
 RT of soluble pyrophosphatase?;  
 RL Microbiology 144:2563-2571(1998).  
 CC -!- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O -> 2 ORTHOPHOSPHATE.

CC -!- COFACTOR: REQUIRES MANGANESE FOR ITS ACTIVITY.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE PPASE CLASS C FAMILY.  
 KW Hydrolase; Manganese.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1828 MW; 3C0E6735D98B38A0 CRC64;

Query Match 18.6%; Score 21; DB 1; Length 16;  
 Best Local Similarity 44.4%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 12 IFATRDQD 20  
 | | :|: |  
 Db 5 IFGHKNPDT 13

RESULT 8  
 APAL\_ERYPA  
 ID APAL\_ERYPA STANDARD; PRT; 20 AA.  
 AC P18647;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE APOLIPOPROTEIN A-I (APO-AI) (FRAGMENT).  
 GN APOA1.  
 OS Erythrocybus patas (Red guenon) (Hussar).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Erythrocybus.  
 OX NCBI\_TaxID=9538;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76184721; PubMed=178359;  
 RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;  
 RT "Characterization of the plasma lipoproteins and apoproteins of the Erythrocybus patas monkey.";  
 RL Biochemistry 15:1928-1933(1976).  
 CC -!- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING THE LECTININ CHOLESTEROL ACYLTRANSFERASE (LCAT).  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -!- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN CHYLOMICRONS.  
 CC -!- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.  
 DR PIR; A05313; A05313.  
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2387 MW; 9C970997C7FC976A CRC64;

Query Match 18.1%; Score 20.5; DB 1; Length 20;  
 Best Local Similarity 36.4%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

QY 2 SPWITDSTDR 12  
 :|: |  
 Db 6 TPW-----DRV 11

RESULT 9  
 GAST\_MACMU  
 ID GAST\_MACMU STANDARD; PRT; 17 AA.  
 AC P33714;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE GASTRIN.  
 GN GAS.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;



CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.  
 CC -!- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M15040; AAA27044.1; -  
 CC DR HSSP; P06612; 1YUA.  
 CC DR StyGene; SG10515; topa.  
 CC DR InterPro; IPR000380; Pro\_topoisomrse.  
 CC DR PROSITE; PS00396; TOPOISOMERASE\_I\_PROK; PARTIAL.  
 CC KW Isomerase; Topoisomerase; DNA-binding.  
 CC FT NON\_TER 1  
 CC SQ SEQUENCE 10 AA; 1145 MW; 2C25B67B02D37338 CRC64;

Query Match 16.8%; Score 19; DB 1; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+03;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 DSPWI 5  
 Db 2 DGKW 6

RESULT 13  
 UHA3\_CANFA STANDARD; PRT; 13 AA.  
 ID UHA3\_CANFA  
 AC P56535;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF HEART TISSUE (SPOT 7520) (FRAGMENT).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 RT dog heart proteins".  
 RL Electrophoresis 18:2795-2802(1997).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC HSC-2DPAGE; P56535; DOG.  
 CC NON\_TER 13  
 FT SEQUENCE 13 AA; 1457 MW; 9F5BD94B9306D76A CRC64;

Query Match 16.8%; Score 19; DB 1; Length 13;  
 Best Local Similarity 33.3%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 WITDSTDRI 12  
 Db 3 FITDPVEXV 11

RESULT 14  
 MY14\_PHEVI STANDARD; PRT; 14 AA.  
 ID MY14\_PHEVI  
 AC P46980;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE MYOACTIVE TETRADECAPEPTIDE (PTP).  
 OS Pheretima vittata (Earthworm).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
 OC Lumbricina; Megascotlecidae; Pheretima.  
 OX NCBI\_TaxID=46674;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Gut;  
 RX MEDLINE=96087879; PubMed=8532604;  
 RA Ukema K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,  
 RA Nomoto K.;  
 RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia  
 RT foetida".  
 RL Peptides 16:995-999(1995).  
 CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT  
 CC MUSCLES.  
 CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.  
 CC Neuropeptide; Amidation.  
 KW MOD\_RES 14 14  
 FT SEQUENCE 14 AA; 1522 MW; DA40BEE67CCD91AD CRC64;

Query Match 16.8%; Score 19; DB 1; Length 14;  
 Best Local Similarity 80.0%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 STDRI 12  
 Db 6 SADRI 10

RESULT 15  
 TKNM\_RANMA STANDARD; PRT; 14 AA.  
 ID TKNM\_RANMA  
 AC P40951;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RANAMARGARIN.  
 OS Rana margaratae (Chinese frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.  
 OX NCBI\_TaxID=121156;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin;  
 RX MEDLINE=90026852; PubMed=2803524;  
 RA Zhu Y.Q., Tian S.H., Wu S.X., Hua J.C., Wu G.F., Zhao E.M., Lu Y.A.,  
 RA Zhu Y.Q., Zou G., Tsou K.;  
 RT "Isolation and structure of ranamargarin, a new tachykinin from the  
 RT skin of Chinese frog Rana margaratae".  
 RL Sci. China, B, Chem. Life Sci. Earth Sci. 32:570-579(1989).  
 RN [2]  
 RP SYNTHESIS.  
 RX MEDLINE=90253600; PubMed=2340087;  
 RA Lu Y.A., Peng J.L., Zhu Y.Q., Wu S.X., Tang Y.Q., Tian S.H., Zou G.;  
 RT "Synthesis and biological activity of a new frog skin peptide,  
 RT ranamargarin".  
 RL Sci. China, B, Chem. Life Sci. Earth Sci. 33:170-177(1990).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC InterPro; IPR002040; Tachykinin.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.  
 FT MOD\_RES 14 14  
 SQ SEQUENCE 14 AA; 1617 MW; D4593AE408C3673D CRC64;

Query Match 16.8%; Score 19; DB 1; Length 14;

Best Local Similarity 60.0%; Pred. No. 2.6e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 DSTDR 11  
|::|  
Db 2 DASDR 6

Search completed: February 21, 2002, 16:52:21  
Job time: 520 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:51:54 ; Search time 60.03 Seconds  
(without alignments)  
51.170 Million cell updates/sec

Title: US-08-753-851-7  
Perfect score: 113  
Sequence: 1 DSPWTDSDRIFATRDQDTI 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 5819

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phage.\*  
10: sp-plant.\*  
11: sp-rodent.\*  
12: sp-virus.\*  
13: sp-vertebrate.\*  
14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	31.9	20	5 P82310	P82310 panulirus j
2	31	27.4	16	4 Q9UC99	Q9UC99 homo sapien
3	30	26.5	18	11 Q62256	Q62256 mus musculu
4	29	25.7	19	11 Q9QV38	Q9QV38 mus sp. erp
5	27	23.9	17	6 Q9TQ28	Q9TQ28 sus scrofa
6	27	23.9	18	4 Q9UGN8	Q9UGN8 homo sapien
7	27	23.9	19	8 Q31687	Q31687 artemia par
8	27	23.9	20	13 Q9PRN6	Q9PRN6 scyllorhinu
9	26	23.0	10	6 Q9N1X1	Q9N1X1 equus cabal
10	26	23.0	15	13 Q9Q594	Q9Q594 gallus gall
11	26	23.0	19	8 Q9T2V5	Q9T2V5 crithidia f
12	26	23.0	20	3 Q9URC7	Q9URC7 saccharomyc
13	26	23.0	21	12 Q83489	Q83489 turnip yell
14	25	22.1	18	6 Q9TRD8	Q9TRD8 oryctolagus
15	25	22.1	19	12 Q91329	Q91329 human immun
16	25	22.1	20	6 Q9N188	Q9N188 equus cabal
17	24	21.2	12	2 Q46747	Q46747 escherichia
18	24	21.2	15	10 Q9S8D5	Q9S8D5 cynara card
19	24	21.2	17	2 P97135	P97135 mycobacteri

Query Match 31.9% ; Score 36; DB 5; Length 20;

20	24	21.2	17	6	O62645	O62645 saguinus oe
21	24	21.2	19	12	Q9VR03	Q9VR03 porcine cir
22	24	21.2	19	12	Q9YIT3	Q9YIT3 bovine circ
23	24	21.2	19	12	O56132	O56132 porcine cir
24	24	21.2	21	11	Q9QV03	Q9QV03 rattus sp.
25	23	20.4	12	3	Q9UR28	Q9UR28 filobasidie
26	23	20.4	15	2	Q9RO21	Q9RO21 salmonella
27	23	20.4	15	4	Q9UCF3	Q9UCF3 homo sapien
28	23	20.4	19	2	Q9FSX2	Q9FSX2 bacillus li
29	23	20.4	19	6	Q9TRH3	Q9TRH3 sus scrofa
30	23	20.4	19	13	Q9PRN1	Q9PRN1 petromyzon
31	23	20.4	20	3	P82262	P82262 aspergillus
32	23	20.4	20	11	Q62546	Q62546 mus spretus
33	23	20.4	21	6	Q9TR10	Q9TR10 bos taurus
34	22.5	19.9	16	11	Q9JHB6	Q9JHB6 mus musculu
35	22	19.5	9	4	Q9UQW0	Q9UQW0 homo sapien
36	22	19.5	9	4	Q9UCQ9	Q9UCQ9 homo sapien
37	22	19.5	13	11	Q63935	Q63935 rattus norv
38	22	19.5	14	11	Q9EPX6	Q9EPX6 mus musculu
39	22	19.5	16	12	Q79459	Q79459 human immun
40	22	19.5	17	4	Q9UM85	Q9UM85 homo sapien
41	22	19.5	17	4	Q9UC87	Q9UC87 homo sapien
42	22	19.5	17	14	Q9EP50	Q9EP50 unidentified
43	22	19.5	18	8	Q9ZV75	Q9ZV75 cenocoelius
44	22	19.5	19	2	Q03977	Q03977 escherichia
45	22	19.5	19	5	Q9TWH2	Q9TWH2 oxytricha g

## ALIGNMENTS

## RESULT 1

P82310 ID P82310 PRELIMINARY; PRT; 20 AA.  
AC P82310:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DE HMOCYANIN IA SUBUNIT (FRAGMENT).  
OS Panulirus japonicus (Japanese spiny lobster).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;  
OC Palinoidea; Palinuridae; Panulirus.  
OX NCBI\_TaxID=6736;  
RN [1]  
RP SEQUENCE  
RC TISSUE=SERUM;  
RX PubMed=3360019;  
RA Makino N., Kimura S.;  
RT "Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.";  
RL Eur. J. Biochem. 173:423-430(1988).

CC -!- FUNCTION: HMOCYANIN ARE COPPER-CONTAINING OXYGEN CARRIERS AND OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND ARTHROPODS.  
CC -!- SUBUNIT: COMPOSED OF 3 MAJOR SUBUNITS (IB, II AND III) AND 1 MINOR SUBUNIT (IA) WHICH FORM HOMOHEXAMERS AND HETEROHEXAMERS. MAY ALSO FORM LARGER STRUCTURES.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HMOCYANIN SUBFAMILY.  
DR InterPro: IPR000896; Hemocyanin.  
DR InterPro: IPR002227; Tyrosinase.  
DR PROSITE; PS00209; HMOCYANIN\_1; PARTIAL.  
DR PROSITE; PS00210; HMOCYANIN\_2; PARTIAL.  
DR PROSITE; PS00497; TYROSINASE\_1; PARTIAL.  
DR PROSITE; PS00498; TYROSINASE\_2; PARTIAL.  
KW Respiratory protein; Oxygen transport; Copper.  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2154 MW; 313BE8E456DDDE09 CRC64;

Best Local Similarity 46.2%; Pred. No. 42;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 DSTDRIFATRDQD 19  
Db 7 DSTDKLLAOKQDD 19

RESULT 2

Q9UC99 PRELIMINARY; PRT; 16 AA.  
AC Q9UC99;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE EOCp-1-EOSINOPHIL-CHEMOTACTIC CYTOKINE.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=95028770; PubMed=7524281;  
RA Schroder J.M., Kameyoshi Y., Christophers E.;  
RT "platelets secrete an eosinophil-chemotactic cytokine which is a  
RT member of the C-C-chemokine family.";  
RL Adv. Exp. Med. Biol. 351:119-128(1993).  
SQ SEQUENCE 16 AA; 1786 MW; 8B041B46AEC2A1 CRC64;

Query Match 27.4%; Score 31; DB 4; Length 16;  
Best Local Similarity 46.2%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SPWITDTRIFA 14  
Db 1 SPYXXDTTPXXFA 13

RESULT 3

Q62256 PRELIMINARY; PRT; 18 AA.  
AC Q62256;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE SPERMATOGONIC-SPECIFIC PROENKEPHALIN.  
GN PENK2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90287163; PubMed=2355920;  
RA Kilpatrick D.L., Zinn S.A., FitzGerald M., Higuchi H., Sabol S.L.,  
RA Meyerhardt J.;  
RT "Transcription of the rat and mouse proenkephalin genes is initiated  
RT at distinct sites in spermatogenic and somatic cells.";  
RL Mol. Cell. Biol. 10:3717-3726(1990).  
DR EMBL: M55181; AAA40127.1; --  
DR MGD; MGI:104628; PenK2.  
SQ SEQUENCE 18 AA; 2043 MW; B96E10CC7049FA76 CRC64;

Query Match 26.5%; Score 30; DB 11; Length 18;  
Best Local Similarity 58.3%; Pred. No. 3.3e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 DSPWITDTRDI 12  
Db 7 DSPW----EDRI 14

RESULT 4  
Q9QV38 PRELIMINARY; PRT; 19 AA.  
AC Q9QV38;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ERP61, GRP58-STRESS-INDUCIBLE LUMINAL ENDOPLASMIC RETICULUM PROTEIN.  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10095;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=94153092; PubMed=8109975;  
RA Mazzarella R.A., Marcus N., Haugejorden S.M., Balcarek J.M.,  
RA Baldassare J.J., Roy B., Li L.J., Lee A.S., Green M.;  
RT "Erp61 is GRP58, a stress-inducible luminal endoplasmic reticulum  
RT protein, but is devoid of phosphatidylinositol-specific phospholipase  
RT C activity.";  
RL Arch. Biochem. Biophys. 308:454-460(1994).  
SQ SEQUENCE 19 AA; 2113 MW; 5A62F136268E3CF4 CRC64;

Query Match 25.7%; Score 29; DB 11; Length 19;  
Best Local Similarity 37.5%; Pred. No. 5.1e+02;  
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 5 ITDSTDTRIFATRDQD 20  
Db 3 VLELTDENFESVSDT 18

RESULT 5

Q9TQZ8 PRELIMINARY; PRT; 17 AA.  
AC Q9TQZ8;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE 20 KDA PHOSPHORYLATION-DEPENDENT PROTEIN PHOSPHATASE-1 INHIBITORY  
DE PROTEIN (FRAGMENT).  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=96362144; PubMed=8720121;  
RA Eto M., Ohmori T., Suzuki M., Furuya K., Morita F.;  
RT "A novel protein phosphatase-1 inhibitory protein potentiated by  
RT protein kinase C. Isolation from porcine aorta media and  
RT characterization.";  
RL J. Biochem. 118:1104-1107(1995).  
SQ SEQUENCE 17 AA; 2109 MW; 472B89943FC5AB15 CRC64;

Query Match 23.9%; Score 27; DB 6; Length 17;  
Best Local Similarity 25.0%; Pred. No. 9.2e+02;  
Matches 4; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 4 WITDSTDTRIFATRDQD 19  
Db 1 WIDGRLELYRGREAD 16

RESULT 6

Q9UGN8 PRELIMINARY; PRT; 18 AA.  
AC Q9UGN8;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)



```

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE APAF1 PROTEIN (FRAGMENT).
GN APAF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Roberts D.L., Dalglish R., Cohen G.M., MacFarlane M.;
RT "The mammalian CED4 homologue, APAF1, exists as two distinct forms in
RT human cells.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ136645; CAB65087.1; -
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 2045 MW; 30D5FA30B885AEF5 CRC64;

Query Match 23.9%; Score 27; DB 4; Length 18;
Best Local Similarity 44.4%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 WITDSTDR1 12
Db 2 WNTDSRSKV 10

RESULT 7
Q31687 ID Q31687 PRELIMINARY; PRT; 19 AA.
AC Q31687;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ATP8E SUBUNIT 8 (FRAGMENT).
GN ATP8.
OS Artemia parthenogenetica.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
OX NCBI_TaxID=6663;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LA MATA;
RX MEDLINE=94223692; PubMed=8169960;
RA Perez M.L., Valverde J.R., Batuecas B., Amat F., Marco R., Garesse R.;
RT "Speciation in the Artemia genus: mitochondrial DNA analysis of
RT bisexual and parthenogenetic brine shrimps.";
RL J. Mol. Evol. 38:156-168(1994).
DR EMBL: X67263; CAA47685.1; -
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 19
FT NON_TER 19
SQ SEQUENCE 19 AA; 2246 MW; 41922AD313E087E3 CRC64;

Query Match 23.9%; Score 27; DB 8; Length 19;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PWIT 6
Db 8 PWIT 11

RESULT 8
Q9PRN6 ID Q9PRN6 PRELIMINARY; PRT; 20 AA.
AC Q9PRN6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

```

```

DE UROTENSIN I HOMOLOG.
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=96051495; PubMed=8536945;
RA Waugh D., Anderson G., Armour K.J., Balment R.J., Hazon N.,
RA Conlon J.M.;
RT "A peptide from the caudal neurosecretory system of the dogfish
RT Scyliorhinus canicula that is structurally related to urotensin I.";
RL Gen. Comp. Endocrinol. 99:333-339(1995).
SQ SEQUENCE 20 AA; 2305 MW; 77A92D52817E97B7 CRC64;

Query Match 23.9%; Score 27; DB 13; Length 20;
Best Local Similarity 46.7%; Pred. No. 1.1e+03;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 7 DSTDRIFATRDQDTI 21
Db 6 NSLDLTFNRRIMDTI 20

RESULT 9
Q9N1X1 ID Q9N1X1 PRELIMINARY; PRT; 10 AA.
AC Q9N1X1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ALCOHOL DEHYDROGENASE 3 (FRAGMENT).
GN ADH3.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20082971; PubMed=10613847;
RA Caetano A.R., Shlue Y.L., Lyons L.A., O'Brien S.J., Laughlin T.F.,
RA Bowling A.T., Murray J.D.;
RT "A comparative gene map of the horse (Equus caballus).";
RL Genome Res. 9:1239-1249(1999).
DR EMBL: AF134056; AAF31299.1; -
FT NON_TER 1
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1144 MW; C4EA25676902D6DD CRC64;

Query Match 23.0%; Score 26; DB 6; Length 10;
Best Local Similarity 75.0%; Pred. No. 7.3e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PWIT 6
Db 7 PWIT 10

RESULT 10
Q90594 ID Q90594 PRELIMINARY; PRT; 15 AA.
AC Q90594;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1999 (TREMBLrel. 09; Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17; Last annotation update)
DE GLOBIN (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.

```

```

OX NCBI_TaxID=9031;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=81261952; PubMed=6266925;
RA Padayatty J., Cummings I., Manske C.L., Higuchi R., Woo S., Salser W.;
RT "Cloning of chicken globin cDNA in bacterial plasmids.";
RL Gene 13:417-422(1981).
CC -!- SIMILARITY: TO GLOBIN FAMILY.
DR EMBL; M10380; AAA48803.1; -.
DR HSSP; P02112; 1HBR.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Respiratory protein.
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 1881 MW; 23E61DA734FA7962 CRC64;

Query Match 23.0%; Score 26; DB 13; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 3 PWITDSTDRIFA 14
Db 8 PW-----TQFFA 15

RESULT 11
Q9T2V5 PRELIMINARY; PRT; 19 AA.
AC Q9T2V5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT 5 (FRAGMENT).
OS Crithidia fasciculata.
OG Mitochondrion.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
OX NCBI_TaxID=5656;
RN [1]
RP SEQUENCE.
RX MEDLINE=97001679; PubMed=8844671;
RA Speijer D., Mulijers A.O., Dekker H., de Haan A., Breek C.K.,
RA Albracht S.P., Benne R.;
RT "Purification and characterization of cytochrome c oxidase from the
RT insect trypanosomatid Crithidia fasciculata.";
RL Mol. Biochem. Parasitol. 79:47-59(1996).
SQ SEQUENCE 19 AA; 2136 MW; 14A22FDB3BD7DD69 CRC64;

Query Match 23.0%; Score 26; DB 8; Length 19;
Best Local Similarity 41.7%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 WITDSTDRIFAT 15
Db 6 WDNASLDTIFSS 17

RESULT 12
Q9URC7 PRELIMINARY; PRT; 20 AA.
AC Q9URC7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE LIPID-BINDING PROTEIN.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
```

```

RX MEDLINE=91353077; PubMed=1882548;
RA Creutz C.E., Snyder S.L., Kambouris N.G.;
RT "Calcium-dependent secretory vesicle-binding and lipid-binding
RT proteins of Saccharomyces cerevisiae.";
RL Yeast 7:229-244(1991).
SQ SEQUENCE 20 AA; 2388 MW; 594377C8C3E72B0D CRC64;

Query Match 23.0%; Score 26; DB 3; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PWITDSTD 10
Db 4 PWDDDET 11

RESULT 13
Q83489 PRELIMINARY; PRT; 21 AA.
AC Q83489;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE (TYPE ISOLATE) COAT PROTEIN (FRAGMENT).
OS Turnip yellow mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.
OX NCBI_TaxID=12154;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88105991; PubMed=3426397;
RA Blok J., Mackenzie A., Guy P., Gibbs A.;
RT "Nucleotide sequence comparisons of turnip yellow mosaic virus
RT isolates from Australia and Europe.";
RL Arch. Virol. 97:283-295(1987).
DR EMBL; M24802; AAA46597.1; -.
DR HSSP; P03608; 1AUY.
KW Coat protein.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2177 MW; FE7DC9B068DD94EA CRC64;

Query Match 23.0%; Score 26; DB 12; Length 21;
Best Local Similarity 62.5%; Pred. No. 1.7e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SPWITDST 9
Db 13 SPLITDTS 20

RESULT 14
Q9TRD8 PRELIMINARY; PRT; 18 AA.
AC Q9TRD8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CHAPERONIN (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=94089752; PubMed=7903455;
RA Rommelaere H., Van Troys M., Gao Y., Melki R., Cowan N.J.,
RA Vandekerckhove J., Ampe C.;
RT "Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and
RT seven related subunits.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:11975-11979(1993).
SQ SEQUENCE 18 AA; 1884 MW; B608F6EBB5A8A2A6 CRC64;
```

Query Match 22.1%; Score 25; DB 6; Length 18;  
Best Local Similarity 35.3%; Pred. No. 2e+03;  
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 5 ITDSTDRI FATRDQDTI 21  
:|| : ||  
Db 2 MTDKGDVTVTNDGATI 18

RESULT 15

O91329 PRELIMINARY; PRT; 19 AA.  
AC O91329;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PATIENT SI FROM FRANCE;  
RX MEDLINE=98285741; PubMed=9621043;  
RA Beltec L., Si.Mohamed A., Muller-trutwin M.C., Gilquin J., Gutmann L.,  
RA Safar M., Barre-Sinoussi F., Kazatchkine M.D.;  
RT "Genetically related human immunodeficiency virus type 1 in three  
adults of a family with no identified risk factor for intrafamilial  
transmission.";  
RL J. Virol. 72:5831-5839(1998).  
DR EMBL; U87220; AAC32980.1; -;  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1  
SQ SEQUENCE 19 AA; 2324 MW; 379CB14A9E073911 CRC64;

Query Match 22.1%; Score 25; DB 12; Length 19;  
Best Local Similarity 44.4%; Pred. No. 2.2e+03;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DSPWITDST 9  
:|| :||  
Db 11 NSTWYSNST 19

Search completed: February 21, 2002, 16:51:55  
Job time: 534 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:42:58 ; Search time 64.88 Seconds  
(without alignments)  
23.976 Million cell updates/sec

Title: US-08-753-851-7

Perfect score: 113

Sequence: 1 DSPWITDSTRIFATRDQDTI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 206698

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	53.1	21	15 AAR53478	CD44 peptide CD44-
2	36	31.9	17	14 AAR33105	Human cytomegalo
3	34	30.1	14	22 AAB64509	Gene 25 human secr
4	34	30.1	15	13 AAR21758	Human FSH agonist
5	34	30.1	15	13 AAR21762	Human FSH agonist
6	34	30.1	15	13 AAR21762	Human FSH agonist
7	33	29.2	14	18 AAR24962	C3/C4/C5-derived c
8	31	27.4	21	10 AAR91156	Artificial peptide
9	30.5	27.0	20	16 AAR74353	HTIV-1 p27(rex) pr
10	30	26.5	8	17 AAR89243	LN clone E13 Vbeta
11	30	26.5	18	18 AAW34412	IL-5 receptor bind

12	30	26.5	18	18 AAW30163	Interleukin-5 bind
13	30	26.5	18	22 AAM15658	Peptide #2092 enco
14	30	26.5	18	22 AAM28159	Peptide #2196 enco
15	30	26.5	18	22 AAM03395	Peptide #2077 enco
16	30	26.5	20	15 AAR45571	Cry j I pollen all
17	30	26.5	20	16 AAR82520	Cry j I Japanese C
18	29.5	26.1	13	21 AAB39650	Anti-IL12 antibody
19	29	25.7	10	18 AAW21661	Peptide useful in
20	29	25.7	10	21 AAY85143	Mouse mammary tumo
21	29	25.7	10	22 AAE05731	Complementarity-de
22	29	25.7	17	22 AAM19218	Peptide #5652 enco
23	29	25.7	17	22 AAM31875	Peptide #5912 enco
24	29	25.7	21	18 AAM16218	Peptide containing
25	29	25.7	21	19 AAW68851	Peptide binding in
26	29	25.7	21	19 AAW58339	IL-1rti binding pe
27	29	25.7	21	20 AAY09987	Interleukin-1 type
28	29	25.7	21	21 AAB17767	IL-1 antagonist pe
29	28.5	25.2	14	13 AAR21764	Human FSH antagoni
30	28	24.8	11	17 AAR86585	Autotaxin peptide
31	28	24.8	13	19 AAW64573	Human Factor V pro
32	28	24.8	16	20 AAY43190	PhoBR pathway apta
33	28	24.8	17	20 AAY02864	Fragment of human
34	28	24.8	19	21 AAB34528	Human secreted pro
35	28	24.8	20	9 AAP80172	Sequence of P85 po
36	28	24.8	20	12 AAR15562	Immunopeptide #2 d
37	28	24.8	20	18 AAW42915	Immunogenic Hepati
38	28	24.8	21	22 AAB69432	Synthetic HAV vp1
39	27	23.9	8	20 AAY10398	T cell epitope/MHC
40	27	23.9	9	17 AAW49397	Human leucocyte an
41	27	23.9	10	22 AAG99393	Proteasome digesti
42	27	23.9	11	22 AAG99391	Proteasome digesti
43	27	23.9	11	22 AAJ00786	Hepatitis C virus
44	27	23.9	11	22 AAJ02408	Hepatitis C virus
45	27	23.9	11	22 AAJ02810	Hepatitis C virus

#### ALIGNMENTS

#### RESULT 1

AAR53478  
ID AAR53478 standard; peptide; 21 AA.

XX AAR53478;

XX 01-DEC-1994 (first entry)

XX CD44 peptide CD44-6.

XX Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;

XX arthritis; tumor cell metastasis; autoimmune disease;

XX immunosuppressive.

XX Homo sapiens.

XX OS

XX WO9409811-A.

XX PD 11-MAY-1994.

XX 29-OCT-1993; 93WO-US10412.

XX 30-OCT-1992; 92US-0973339.

XX (UYDU-) UNIV DUKE.

XX Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;

XX WPI; 1994-167121/20.

XX Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis

PS Claim 4; Page 14; 83pp; English.

CC The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 200-219 of the CD44 protein sequence.

CC Sequence 21 AA;

Query Match 53.1%; Score 60; DB 15; Length 21;  
Best Local Similarity 91.7%; Pred. No. 0.0062;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DSPWITDSTDRI 12  
          ||||||| II

DB 10 dspwldstprl 21

RESULT 2

AAR33105  
ID AAR33105 standard; peptide; 17 AA.

XX AAR33105;

XX 29-JUN-1993 (first entry)

DE Human cytomegalovirus antibody specific peptide 3.2.

KW hCMV; CMV; detection; antibody specificity determination.

XX Synthetic.

XX DE4128684-A.

XX 04-MAR-1993.

XX 29-AUG-1991; 91DE-4128684.

XX 29-AUG-1991; 91DE-4128684.

XX (BEHW ) BEHRINGWERKE AG.

PI Stuber W, Wiczorek L, Ziegelmaier R;

XX WPI; 1993-077562/10.

XX New peptide(s) specific for human cytomegalovirus - useful to detect antibodies specific for the specified virus, and for preparing vaccines

PS Claim 2; Page 24; 25pp; German.

XX The sequence is that of a peptide which specifically reacts with antibodies against human cytomegalovirus (HCMV). It can be used as part of immunochemical methods for the detection of HCMV antibodies. It can also be used, with at least one immunoreactive peptide from a pathogen other than HCMV, in an immunological combination process for the determination of different antibody specificities of different pathogens. The other pathogen is pref. HSV-1 or -2, EBV, VZV, HHV8, HAV, HCV, and/or HIV-1 or -2. Use of the peptide avoids the false signals of prior art prods. resulting from cross-contamination during the purification of HCMV antigen from cell cultures.

XX Sequence 17 AA;

Query Match 31.9%; Score 36; DB 14; Length 17;  
Best Local Similarity 50.0%; Pred. No. 31;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 7 DSTDRIFATRDQ 18  
          | : : | |||  
DB 4 daadevwalrdq 15

RESULT 3

AAB64509

ID AAB64509 standard; Protein; 14 AA.

XX AC AAB64509;

XX 23-MAR-2001 (first entry)

DE Gene 25 human secreted protein homologous amino acid sequence #147.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; autoimmune disease; rheumatoid arthritis; hyperproliferative disorder; neoplasm; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection; angiogenesis; nervous system disorder; Alzheimer's disease; skin aging; ocular disorder; corneal infection; wound healing; food additive; preservative.

XX Homo sapiens.

OS WO200077255-A1.

XX 21-DEC-2000.

XX 01-JUN-2000; 2000WO-US14926.

XX 11-JUN-1999; 99US-0138628.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis GA;

XX WPI; 2001-025337/03.

XX Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -

PS Disclosure; Page 564; 593pp; English.

XX The polynucleotide sequences given in AAF32699 to AAF32747 encode the human secreted proteins given in AAB64422 to AAB64470. AAB64471 to AAB64548 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: antiarthritic; immunosuppressive; antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The polynucleotides and polypeptides can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. AAF32690 to AAF32698 and AAB64421 represent sequences used in the exemplification of the present invention.

XX Sequence 14 AA;

Query Match 30.1%; Score 34; DB 22; Length 14;  
 Best Local Similarity 60.0%; Pred. No. 52;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 8 STDTRIFATRD 17  
 ||| | : | : |  
 Db 3 stprvfttgd 12

RESULT 4  
 AAR21758  
 ID AAR21758 standard; Protein; 15 AA.  
 XX AC AAR21758;  
 XX DT 23-JUN-1992 (first entry)  
 XX DE Human FSH agonist peptide #4.  
 XX KW Follicle stimulating hormone; FSH; thyroid stimulating hormone; TSH;  
 XX KW contraceptive; in vitro fertilisation.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Region 1..5  
 XX FT /note= "amino acids 92, 88, 89, 90 and 91 from  
 XX FT beta-chain of human FSH"  
 XX FT Modified-site 6  
 XX FT /label= Acp  
 XX FT Region 7..10  
 XX FT /note= "partial peptide from alpha-chain of  
 XX FT glycopeptide hormones"  
 XX FT Modified-site 11  
 XX FT /label= OTHER  
 XX FT /note= "5-aminopentanoic acid"  
 XX FT Region 12..15  
 XX FT /note= "amino acids 34-37 of beta-chain of hFSH"  
 XX WO9202542-A.  
 XX PN 20-FEB-1992.  
 XX PD 26-JUL-1991; 91WO-NL00139.  
 XX PR 27-JUL-1990; 90NL-0001709.  
 XX PA (DIER-) STICHT CENT DIERGEN.  
 XX PI Hage Van Noort M, Puijk WC, Meloen RH;  
 XX DR WPI; 1992-080024/10.  
 XX PS New peptide having glyco-protein hormone agonist or antagonist  
 XX PT action - useful as vaccines for passive immunisation against  
 XX PT glyco-protein hormone and to regulate fertility  
 XX PS Claim 9; Page 74; 82pp; English.  
 XX CC The invention also covers substitution, deletion and insertion  
 XX CC variants having FSH agonistic activity, as well as derivs. in which  
 XX CC the free amino group of the N-terminal amino acid and/or the free  
 XX CC carboxyl group of the C-terminal amino acid are blocked or otherwise  
 XX CC modified. The agonist can be used, e.g. to improve sperm production  
 XX CC in agricultural domestic animals such as horses and bulls.  
 XX CC See AAR21754-R21765 and AAR23788-9.  
 XX Sequence 15 AA;

Query Match 30.1%; Score 34; DB 13; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 56;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 TDSTDTRIFATRD 17  
 :||| | : |||  
 Db 3 sdxsrayxttd 14

RESULT 5  
 AAR21762  
 ID AAR21762 standard; Protein; 15 AA.  
 XX AC AAR21762;  
 XX DT 23-JUN-1992 (first entry)  
 XX DE Human FSH antagonist peptide #4.  
 XX KW Follicle stimulating hormone; FSH; thyroid stimulating hormone; TSH;  
 XX KW contraceptive; in vitro fertilisation.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Region 1..5  
 XX FT /note= "amino acids 92, 88, 89, 90 and 91 from  
 XX FT beta-chain of human FSH"  
 XX FT Modified-site 6  
 XX FT /label= Acp  
 XX FT Region 7..10  
 XX FT /note= "partial peptide from alpha-chain of  
 XX FT glycopeptide hormones"  
 XX FT Region 12..15  
 XX FT /note= "amino acids 34-37 of beta-chain of hFSH"  
 XX WO9202542-A.  
 XX PN 20-FEB-1992.  
 XX PD 26-JUL-1991; 91WO-NL00139.  
 XX PR 27-JUL-1990; 90NL-0001709.  
 XX PA (DIER-) STICHT CENT DIERGEN.  
 XX PI Hage Van Noort M, Puijk WC, Meloen RH;  
 XX DR WPI; 1992-080024/10.  
 XX PS New peptide having glyco-protein hormone agonist or antagonist  
 XX PT action - useful as vaccines for passive immunisation against  
 XX PT glyco-protein hormone and to regulate fertility  
 XX PS Claim 10; Page 74; 82pp; English.  
 XX CC The invention also covers substitution, deletion and insertion  
 XX CC variants having FSH antagonistic activity, as well as derivs. in  
 XX CC which the free amino group of the N-terminal amino acid and/or  
 XX CC the free carboxyl group of the C-terminal amino acid are blocked  
 XX CC or otherwise modified. The antagonist can be used, e.g. as an  
 XX CC alternative "steroid-free" pill to block ovulation or as an  
 XX CC immunogen for active or passive immunisation against gonadotropins.  
 XX CC See AAR21754-R21765 and AAR23788-9.  
 XX Sequence 15 AA;

Query Match 30.1%; Score 34; DB 13; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 56;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 TDSTDRIFATRD 17  
:||| : |||  
Db 3 sdsxraygtrd 14

RESULT 6  
AAR21763  
ID AAR21763 standard; Protein; 15 AA.  
XX  
AC AAR21763;

XX 23-JUN-1992 (first entry)

XX Human FSH antagonist peptide #5.

XX Follicle stimulating hormone; FSH; thyroid stimulating hormone; TSH;  
KW contraceptive; in vitro fertilisation.

XX Synthetic.

XX Key Location/Qualifiers  
FH Region 1..5  
FT /note= "amino acids 92, 88, 89, 90 and 91 from  
FT beta-chain of human FSH"  
FT Modified-site 6  
FT /label= Acp  
FT Region 7..10  
FT /note= "partial peptide from alpha-chain of  
FT glycopeptide hormones"  
FT Modified-site 11  
FT /label= bala  
FT Region 12..15  
FT /note= "amino acids 34-37 of beta-chain of hFSH"

XX WO9202542-A.

XX 20-FEB-1992.

XX 26-JUL-1991; 91WO-NL00139.

XX 27-JUL-1990; 90NL-0001709.

XX (DIER-) STICHT CENT DIERGEN.

XX Hage Van Noort M, Puijck WC, Meloen RH;

XX WPI; 1992-080024/10.

XX New peptide having glyco-protein hormone agonist or antagonist  
PT action - useful as vaccines for passive immunisation against  
PT glyco-protein hormone and to regulate fertility

XX Claim 10; Page 75; 82pp; English.

XX The invention also covers substitution, deletion and insertion  
CC variants having FSH antagonistic activity, as well as derivs. in  
CC which the free amino group of the N-terminal amino acid and/or  
CC the free carboxyl group of the C-terminal amino acid are blocked  
CC or otherwise modified. The antagonist can be used, e.g. as an  
CC alternative "steroid-free" pill to block ovulation or as an  
CC immunogen for active or passive immunisation against gonadotropins.  
CC See AAR21754-R21765 and AAR23788-9.

XX Sequence 15 AA;

Query Match 30.1%; Score 34; DB 13; Length 15;  
Best Local Similarity 50.0%; Pred. No. 56;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 TDSTDRIFATRD 17  
:||| : |||  
Db 3 sdsxraygtrd 14

RESULT 7

AAW24962  
ID AAW24962 standard; peptide; 14 AA.

XX AAW24962;

XX 28-OCT-1997 (first entry)

XX C3/C4/C5-derived complement inhibitory peptide C4-B1.

XX Indel; insertion; deletion; protein family; sequence alignment;  
KW interaction site; ligand; interface peptide; binding complex;  
KW modulation; complement; vaccine; antibody; tissue rejection;  
KW ischaemia; aneurysm; chronic inflammation; autoimmune disease.

XX Synthetic.

XX Key Location/Qualifiers  
FH Modified-site 1  
FT /note= "acylated N-terminus"

XX WO9701578-A1.

XX 16-JAN-1997.

XX 27-JUN-1996; 96WO-US10958.

XX 14-JUN-1996; 96US-0663617.

XX 29-JUN-1995; 95US-0000674.

XX (MED1-) MEDICAL BIOLOGY INST.

XX Ogata RT;

XX WPI; 1997-100166/09.

XX Modulatory cpds. identified by screening peptide(s) derived from  
PT indels - esp. from regions involved in protein-protein interaction,  
PT useful as inhibitors or activators of the complement system

XX Claim 30; Page 23; 161pp; English.

XX The peptides AAW24940-69 represent peptides derived from "indel"  
CC sequences. Indels are peptide sequences corresponding to insertion  
CC or deletion sequences in members of a family of proteins when their  
CC sequences are aligned and compared. The peptides often correspond  
CC to sites of interaction between proteins or proteins and their ligands.  
CC The invention relates to a method for identifying regions of contact  
CC (interface peptides) in protein-protein binding complexes and for the  
CC preparation of modulators of this interaction. The peptides presented  
CC here are modulators of the complement system (including peptides that  
CC act as vaccines by inducing modulating antibodies) which are useful  
CC therapeutically, e.g. to prevent or treat tissue rejection, ischaemia,  
CC aneurysm, chronic inflammation, autoimmune diseases, etc. This peptide  
CC is derived from the C3/C4/C5 family of proteins in a region proximal to  
CC indel-2.

XX Sequence 14 AA;

Query Match 29.2%; Score 33; DB 18; Length 14;  
Best Local Similarity 71.4%; Pred. No. 74;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SPWITDS 8

Db 8 spwikds 14

RESULT 8

AAP91156

ID AAP91156 standard; peptide; 21 AA.  
 XX  
 AC AAP91156;  
 XX  
 DT 26-APR-1990 (first entry)  
 XX  
 DE Artificial peptide containing a sequence which comprises an epitope  
 DE of HIV.  
 XX  
 XX Artificial HIV peptide; HIV epitope; immunoassay kit; HIV vaccine;  
 KW artificial antigen.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 5..11  
 FT /note="This sequence is specifically claimed."  
 XX  
 PN WO8903844-A.  
 XX  
 XX 05-MAY-1989.  
 XX  
 XX 27-OCT-1988; 88WO-SE00570.  
 XX  
 PR 28-OCT-1987; 87SE-0004185.  
 XX  
 PA (FERR ) FERRING AB.  
 XX  
 PI Trojnar J, Wahren B, Ruden U;  
 XX  
 DR WPI; 1989-150751/20.  
 XX  
 PT Artificial HIV peptide  
 PT - having a sulphur bridge between 2 cysteine residues  
 PT located on each side of a HIV epitope  
 XX  
 PS Claim 2; Page 21, lines 11-13; 27pp; English.  
 XX  
 CC It has an amino acid sequence which corresp. to a naturally occurring  
 CC amino acid sequence for an epitope of HIV, and which further has two  
 CC Cys residues on each side of the epitope. It is stabilised by a sulphur  
 CC bridge between the 2 Cys residues formed by a chemical oxidation step.  
 CC Also claimed are peptides having a shorter sequence. It provides  
 CC an assay for the detn. of antibodies induced by HIV for use in  
 CC diagnostic immunoassay kits. It may also be used as an immunising  
 CC component in vaccine compsns. against HIV.  
 XX  
 XX Sequence 21 AA;  
 SQ  
 Query Match 27.4%; Score 31; DB 10; Length 21;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 PWITDS 8  
 DB 16 pwnds 21  
 RESULT 9  
 ID AAR74353  
 XX AAR74353 standard; peptide; 20 AA.  
 XX  
 AC AAR74353;  
 XX  
 DT 23-JAN-1996 (first entry)  
 XX  
 DE HTLV-1 p27(rex) protein antigenic peptide rex2.  
 XX  
 KW Human T-cell lymphotropic virus type I; HTLV-I; p27(rex) protein;  
 KW antigenic peptide; rex2; residues 16-35; adult T-cell leukaemia;  
 KW HTLV-I associated myelopathy.  
 XX  
 OS Human lymphotropic virus type I.  
 XX

PN US5420244-A.  
 XX  
 PD 30-MAY-1995.  
 XX  
 PF 06-AUG-1993; 93US-0103742.  
 XX  
 PR 06-AUG-1993; 93US-0103742.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Lal RB, Rudolph DL;  
 PI WPI; 1995-206307/27.  
 XX  
 DR New HTLV-I antigenic peptide(s) - used for diagnosing HTLV-I  
 PT infection(s), partic. HTLV-I associated myelopathy and adult T-cell  
 PT leukemia  
 XX  
 PS Claim 6; Columns 11-12; 21pp; English.  
 XX  
 CC AAR74352-R74355 are claimed human T-cell lymphotropic virus type I  
 CC (HTLV-I) p27(rex) protein antigenic peptides. They can be used to  
 CC diagnose diseases caused by HTLV-I, e.g. adult T-cell leukaemia,  
 CC and HTLV-I associated myelopathy.  
 XX  
 XX Sequence 20 AA;  
 SQ  
 Query Match 27.0%; Score 30.5; DB 16; Length 20;  
 Best Local Similarity 38.9%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 3; Mismatches 7; Indels 1; Gaps 1;  
 QY 2 SPWIT-DSTDRIFATRDQ 18  
 DB 3 tpwptsggldrvffsdgtg 20  
 RESULT 10  
 ID AAR89243  
 XX AAR89243 standard; peptide; 8 AA.  
 AC AAR89243;  
 XX  
 DT 05-SEP-1996 (first entry)  
 XX  
 DE LN clone E13 Vbeta8-CDR3.  
 XX  
 KW Polymerase chain reaction; PCR; primer; amplify; human; T cell receptor;  
 KW beta chain; TCR; myelin basic protein; BP; autoantigen; encephalitogen;  
 KW experimental autoimmune encephalomyelitis; EAE; multiple sclerosis; MS;  
 KW autoimmune disease; neurological disease; cerebrospinal fluid; therapy;  
 KW central nervous system; complementarity determining region; CDR;  
 KW T lymphocyte; optical nerve damage; anterior chamber inflammation.  
 XX  
 OS Synthetic.  
 XX  
 XX WO9601329-A1.  
 XX  
 PD 18-JAN-1996.  
 XX  
 PF 26-JUN-1995; 95WO-US08086.  
 XX  
 PR 01-JUL-1994; 94US-0270634.  
 XX  
 PA (CONN-) CONNECTIVE THERAPEUTICS INC.  
 XX  
 PI Buenafe A, Offner H, Vandenbark AA;  
 XX  
 DR WPI; 1996-087679/09.  
 DR N-PSDB; AAT10620.  
 XX  
 PT Methods for diagnosis and immune-related therapy of autoimmune  
 PT diseases - partic. multiple sclerosis, by detecting marker T cell





sequences a molecular weight below 5 kD, and an IC50 of 100 microm or less for binding to IL5 receptors. The peptide may have one or more peptide CONH linkages replaced, an altered N-terminal group, and a C-terminal group of formula COR2, where R2 = OH, lower alkoxy or NR3R4, where R3 and R4 = H or lower alkyl, or the nitrogen atom of the -NR3R4 group can optionally be the amine group of the N-terminus of the peptide so as to form a cyclic peptide. The method may be used for treating an IL-5-mediated disorder that involves the accumulation of eosinophils. It is especially for treating an IL-5-mediated inflammatory disorder of the respiratory tract, especially asthma, by administering an inhalant composition comprising a peptide or peptide analogue as above in an aerosolised carrier solution or dry powder. Other diseases that may be treated with the method include rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis, inflammatory bowel disease, and various skin conditions (including urticaria, angioedema, eczematous dermatitis and psoriasis).

Sequence 18 AA;

Query Match 26.5%; Score 30; DB 18; Length 18;  
Best Local Similarity 53.8%; Pred. No. 2.9e+02;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 7 DSTDRIFATRDOD 19  
| | | | | |  
Db 2 dgctrivatrswd 14

RESULT 13

AAM15658  
ID AAM15658 standard; Protein; 18 AA.

XX AAM15658;

DT 12-OCT-2001 (first entry)

DE Peptide #2092 encoded by probe for measuring cervical gene expression.

XX Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.

XX Homo sapiens.

XX WO200157278-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -

XX Claim 27; SEQ ID No 20484; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for

CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 18 AA;

Query Match 26.5%; Score 30; DB 22; Length 18;  
Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SPWITDST 9  
| | | | |  
Db 10 spwcpnst 17

RESULT 14

AAM28159  
ID AAM28159 standard; Protein; 18 AA.

XX AAM28159;

XX 17-OCT-2001 (first entry)

DE Peptide #2196 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488997/53.

XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
PS Claim 27; SEQ ID No 28428; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs:  
CC see AAI31315-AA157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX Sequence 18 AA;

Query Match 26.5%; Score 30; DB 22; Length 18;  
Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Search completed: February 21, 2002, 16:42:58  
Job time: 232 sec

QY 2 SPWITDST 9  
Db 10 spwcpnst 17  
||| :||  
10 spwcpnst 17

RESULT 15  
AAM03395  
ID AAM03395 standard; Protein: 18 AA.  
XX  
AC AAM03395;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #2077 encoded by probe for measuring breast gene expression.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
PS Claim 27; SEQ ID No 12135; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes  
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 18 AA;

Query Match 26.5%; Score 30; DB 22; Length 18;  
Best Local Similarity 62.5%; Pred. NO. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SPWITDST 9  
Db 10 spwcpnst 17  
||| :||  
10 spwcpnst 17



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:43:36 ; Search time 31.57 Seconds  
(without alignments)  
14.969 Million cell updates/sec

Title: US-08-753-851-7  
Perfect score: 113  
Sequence: 1 DSPWIDSDRIFATRDQDTI 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 114342

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	87.6	20	US-08-143-311B-7	Sequence 7, Appl
2	60	53.1	21	US-08-143-311B-6	Sequence 6, Appl
3	36	31.9	17	US-08-460-874A-31	Sequence 31, Appl
4	36	31.9	17	US-08-388-883B-31	Sequence 31, Appl
5	36	31.9	17	US-08-462-211A-31	Sequence 31, Appl
6	35	31.0	15	US-08-221-583-31	Sequence 31, Appl
7	35	31.0	15	PCT-US95-04018-31	Sequence 31, Appl
8	33	29.2	15	US-08-221-583-32	Sequence 32, Appl
9	33	29.2	15	PCT-US95-04018-32	Sequence 32, Appl
10	30.5	27.0	20	US-08-103-742-6	Sequence 6, Appl
11	30	26.5	18	US-08-478-312-13	Sequence 13, Appl
12	30	26.5	18	US-08-485-302-13	Sequence 13, Appl
13	30	26.5	20	US-08-467-023-55	Sequence 55, Appl
14	29	25.7	10	US-08-555-394-16	Sequence 16, Appl
15	29	25.7	10	US-08-745-892-16	Sequence 16, Appl
16	29	25.7	10	US-09-085-072-3	Sequence 3, Appl
17	29	25.7	15	US-08-937-102-21	Sequence 21, Appl
18	29	25.7	15	US-08-937-102-22	Sequence 22, Appl
19	29	25.7	15	US-08-937-102-23	Sequence 23, Appl
20	29	25.7	21	US-08-190-788A-251	Sequence 251, App
21	29	25.7	21	US-08-383-474B-254	Sequence 254, App
22	29	25.7	21	US-08-465-391A-251	Sequence 251, App
23	29	25.7	21	US-08-464-538B-251	Sequence 251, App
24	29	25.7	21	US-08-463-076E-308	Sequence 308, App
25	28	24.8	9	US-08-463-076E-308	Patent No. 5217891
26	28	24.8	11	US-08-346-455B-43	Sequence 43, Appl
27	28	24.8	11	US-08-977-221-43	Sequence 43, Appl

28 28 24.8 11 5 PCT-US95-06613-43 Sequence 43, Appl  
29 28 24.8 20 2 US-08-934-915-40 Sequence 40, Appl  
30 27 23.9 14 1 US-07-745-382-7 Sequence 7, Appl  
31 27 23.9 14 1 US-07-921-848-7 Sequence 7, Appl  
32 27 23.9 14 1 US-08-165-301A-7 Sequence 7, Appl  
33 27 23.9 14 4 US-08-810-436-7 Sequence 7, Appl  
34 27 23.9 14 5 PCT-US94-14179-7 Sequence 7, Appl  
35 27 23.9 15 1 US-08-221-583-8 Sequence 8, Appl  
36 27 23.9 15 2 US-08-553-257A-49 Sequence 49, Appl  
37 27 23.9 15 5 PCT-US95-04018-8 Sequence 8, Appl  
38 27 23.9 17 1 US-08-451-947-32 Sequence 32, Appl  
39 27 23.9 17 2 US-08-424-826A-32 Sequence 32, Appl  
40 27 23.9 17 3 US-08-928-694-32 Sequence 32, Appl  
41 27 23.9 17 5 PCT-US91-06950-32 Sequence 32, Appl  
42 27 23.9 18 2 US-08-572-951-11 Sequence 11, Appl  
43 27 23.9 18 2 US-08-572-951-29 Sequence 29, Appl  
44 27 23.9 19 3 US-08-851-843A-185 Sequence 185, App  
45 27 23.9 19 4 US-08-974-549A-304 Sequence 304, App

## ALIGNMENTS

RESULT 1  
US-08-143-311B-7  
; Sequence 7, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEEBOE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-7

Query Match 87.6%; Score 99; DB 2; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.6e-10;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSPWITDSTRIPATRDQDT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 DSPWITDSTRIPATRDQDT 20

RESULT 2  
US-08-143-311B-6  
; Sequence 6, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-6

Query Match 53.1%; Score 60; DB 2; Length 21;  
Best Local Similarity 91.7%; Pred. No. 0.0014;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSPWITDSTRIP 12  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 DSPWITDSTRIP 21

RESULT 3

US-08-460-874A-31  
; Sequence 31, Application US/08460874A  
; Patent No. 5744298  
; GENERAL INFORMATION:  
; APPLICANT: Stuber, Werner  
; APPLICANT: Wiczorek, Leszek  
; APPLICANT: Ziegelmair, Robert  
; TITLE OF INVENTION: HCMV-Specific Peptides, Agents Therefor  
; TITLE OF INVENTION: and the Use Thereof  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington,  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,874A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/936,219  
; FILING DATE: 27-AUG-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P4128684.7  
; FILING DATE: 29-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Forman, David S.  
; REGISTRATION NUMBER: 33,694  
; REFERENCE/DOCKET NUMBER: 05552-1210-04000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-460-874A-31

Query Match 31.9%; Score 36; DB 1; Length 17;  
Best Local Similarity 50.0%; Pred. No. 7.9;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 DSTDRIFATRDQ 18  
| : | : | | | | | | | | | | | | | | | | | | | | | |  
Db 4 DAADEVWALRDQ 15

RESULT 4  
US-08-388-883B-31  
; Sequence 31, Application US/08388883B  
; Patent No. 5859185  
; GENERAL INFORMATION:  
; APPLICANT: ST BER, Werner  
; APPLICANT: WICZOREK, Leszek  
; APPLICANT: ZIEGELMAIER, Robert  
; TITLE OF INVENTION: HCMV-Specific Peptides, Agents Therefor  
; TITLE OF INVENTION: and the Use Thereof  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner L.L.P.

```

; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington,
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388.883B
; FILING DATE: 13-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300.305
; FILING DATE: 23-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/936.219
; FILING DATE: 27-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P4128684.7
; FILING DATE: 29-AUG-1991
; NAME: Forman, David S.
; REGISTRATION NUMBER: 33,694
; REFERENCE/DOCKET NUMBER: 5552-1210-02000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-388-883B-31

Query Match 31.9%; Score 36; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 DSTDRIFATRDQ 18
Db 4 DAADEVWALRDQ 15

RESULT 5
US-08-462-211A-31
; Sequence 31, Application US/08462211A
; Patent No. 6143493
; GENERAL INFORMATION:
; APPLICANT: Stuber, Werner
; APPLICANT: Wiczorek, Leszek
; APPLICANT: Ziegelmaler, Robert
; TITLE OF INVENTION: HCMV-Specific Peptides, Agents Therefor
; TITLE OF INVENTION: and the Use Thereof
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington,
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/462,211A
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/388,883
; FILING DATE: 13-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,305
; FILING DATE: 23-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/936,219
; FILING DATE: 27-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P4128684.7
; FILING DATE: 29-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Forman, David S.
; REGISTRATION NUMBER: 33,694
; REFERENCE/DOCKET NUMBER: 5552.1210-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-462-211A-31

Query Match 31.9%; Score 36; DB 4; Length 17;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 DSTDRIFATRDQ 18
Db 4 DAADEVWALRDQ 15

RESULT 6
US-08-221-583-31
; Sequence 31, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids

```

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-31

Query Match 31.0%; Score 35; DB 1; Length 15;  
Best Local Similarity 45.5%; Pred. No. 9.9;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 11 RIFATRDQDTI 21  
Db 3 KVFAKTSDTV 13

## RESULT 7

PCT-US95-04018-31  
; Sequence 31, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-31

Query Match 31.0%; Score 35; DB 5; Length 15;  
Best Local Similarity 45.5%; Pred. No. 9.9;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 11 RIFATRDQDTI 21  
; :|||: ||:

Db 3 KVFAKTSDTV 13

## RESULT 8

US-08-221-583-32  
; Sequence 32, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-32

Query Match 29.2%; Score 33; DB 1; Length 15;  
Best Local Similarity 50.0%; Pred. No. 21;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 12 IFATRDQDTI 21  
Db 1 VFATKTSDTV 10

## RESULT 9

PCT-US95-04018-32  
; Sequence 32, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04018  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,580  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,583  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,581  
FILING DATE: 01-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0232  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-04018-32

Query Match 29.2%; Score 33; DB 5; Length 15;  
Best Local Similarity 50.0%; Pred. No. 21;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 12 IFATRDQDTI 21  
DB 1 VFATKSTDV 10

RESULT 10  
US-08-103-742-6  
Sequence 6, Application 08/103742  
Patent No. 5420244  
GENERAL INFORMATION:  
APPLICANT: RUDOLPH, DONNA L.  
APPLICANT: LAL, RENU B.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR  
DIAGNOSING HTLV-I ASSOCIATED MYELOPATHY AND ADULT T-CELL LEUKEMIA  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, 127 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 08/103,742  
FILING DATE: 06 AUG 1993  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: SPRATT, GWENDOLYN D.  
REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 1414.080  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-103-742-6

Query Match 27.0%; Score 30.5; DB 1; Length 20;  
Best Local Similarity 38.9%; Pred. No. 74;  
Matches 7; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 2 SPWIT-DSTDRIFATRDQ 18  
DB 3 TPWTSQGLDRVFSDTQ 20

RESULT 11  
US-08-478-312-13  
Sequence 13, Application US/08478312  
Patent No. 5654276  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: England, Bruce  
APPLICANT: Schatz, Peter  
APPLICANT: Sloan, Derek  
APPLICANT: Chen, Min-Jia  
TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5  
NUMBER OF SEQUENCES: 59  
NUMBER OF INVENTION: Receptor  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, N.V.  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,312  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1088.1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-478-312-13

Query Match 26.5%; Score 30; DB 1; Length 18;  
Best Local Similarity 53.8%; Pred. No. 79;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 7 DSTDRIFATRDQD 19  
DB 2 DGCTRIVATRSWD 14

```
RESULT 12
US-08-485-302-13
; Sequence 13, Application US/08485302
; Patent No. 5668110
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: England, Bruce
; APPLICANT: Schatz, Peter
; APPLICANT: Sloan, Min-Jia
; APPLICANT: Chen, Min-Jia
; TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, N.V.
; STREET: 4001 Miranda Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,302
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1088.1b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-485-302-13

Query Match 26.5%; Score 30; DB 1; Length 18;
Best Local Similarity 53.8%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 7 DSTDRIFATRDQD 19
DB 2 DGTCTIVATRSWD 14

RESULT 13
US-08-467-023-55
; Sequence 55, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.;
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garman, Richard D;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Yeung, Siu-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; TITLE OF INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-467-023-55

Query Match 26.5%; Score 30; DB 3; Length 20;
Best Local Similarity 41.7%; Pred. No. 89;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SPWITDSTDRIFF 13
DB 2 SNWVWQSTQDVF 13

RESULT 14
US-08-555-394-16
; Sequence 16, Application US/08555394
; Patent No. 5686247
; GENERAL INFORMATION:
; APPLICANT: HOLLAND, JAMES
; APPLICANT: POGO, BEATRIZ
; TITLE OF INVENTION: DETECTION OF MAMMARY TUMOR VIRUS ENV GENE-
; TITLE OF INVENTION: LIKE SEQUENCES IN HUMAN BREAST CANCER
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,394
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kole, Lisa B  
; REGISTRATION NUMBER: 35,225  
; REFERENCE/DOCKET NUMBER: A30363 - 165/32494  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-408-2628  
; TELEFAX: 212-765-2519  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
US-08-553-394-16

Query Match 25.7%; Score 29; DB 1; Length 10;  
Best Local Similarity 57.1%; Pred. No. 56;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSPWITD 7  
| | | |  
Db 4 DEPWDD 10

RESULT 15  
US-08-745-892-16  
; Sequence 16, Application US/08745892  
; Patent No. 6040146  
; GENERAL INFORMATION:  
; APPLICANT: HOLLAND, JAMES  
; APPLICANT: POGO, BEATRIZ  
; TITLE OF INVENTION: DETECTION OF MAMMARY TUMOR VIRUS-LIKE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10112-0228  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/745,892  
; FILING DATE: 08-NOV-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/555,394  
; FILING DATE: 09-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kole, Lisa B  
; REGISTRATION NUMBER: 35,225  
; REFERENCE/DOCKET NUMBER: A30363-A - 165/35211  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-408-2628  
; TELEFAX: 212-765-2519  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
US-08-745-892-16

Query Match 25.7%; Score 29; DB 3; Length 10;  
Best Local Similarity 57.1%; Pred. No. 56;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DSPWITD 7  
| | | |  
Db 4 DEPWDD 10

Search completed: February 21, 2002, 16:43:37  
Job time: 171 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:50:51 ; Search time 13.27 Seconds  
(without alignments)  
200.913 Million cell updates/sec

Title: US-08-753-851-8  
Perfect score: 200  
Sequence: 1 CATRDQDTFHPSGSHTHESEDHSHGSGEGGAN 35

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 7547

Minimum DB seq length: 0  
Maximum DB seq length: 35

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	31.5	29	A41683	hyaluronate recept
2	42	21.0	34	C31514	hemopexin - chicke
3	38	19.0	19	PQ0548	capsid protein VP2
4	35	17.5	28	PQ0263	dnak-type molecula
5	34.5	17.2	26	T04371	thumatin-like pro
6	33	16.5	15	I29501	fibrinopeptide A -
7	33	16.5	24	F45357	Kex2/subtilisin-li
8	33	16.5	24	C45357	Kex2/subtilisin-li
9	33	16.5	27	P23422	second envelope g1
10	33	16.5	31	I52332	tau protein - huma
11	33	16.5	34	E45357	Kex2/subtilisin-li
12	33	16.5	34	B45357	Kex2/subtilisin-li
13	32	16.0	18	B32473	histidine-rich pro
14	32	16.0	31	B82151	hypothetical prote
15	31	15.5	15	JP0101	fibrinogen alpha c
16	31	15.5	22	B81078	hypothetical prote
17	31	15.5	31	B82138	hypothetical prote
18	30	15.0	21	S39455	34K ribonucleot
19	30	15.0	27	P23443	second envelope g1
20	30	15.0	27	P23444	second envelope g1
21	30	15.0	28	I59477	antigen, T-cell re
22	30	15.0	34	S13439	lectin II, anti-H
23	30	15.0	35	E44957	chitinase (EC 3.2.
24	30	15.0	35	T42121	hypothetical prote
25	29	14.5	15	S57577	T cell receptor v-
26	29	14.5	17	S69164	ferredoxin al - Ja
27	29	14.5	21	S48632	carbonic anhydrase
28	29	14.5	27	P23336	second envelope g1
29	29	14.5	27	P23335	second envelope g1

30 29 14.5 27 2 PC2334 second envelope g1  
31 29 14.5 27 2 S57798 PK3 protein - toma  
32 29 14.5 28 2 A60359 pollen allergen DG  
33 29 14.5 28 2 A60692 proline-rich prote  
34 29 14.5 33 2 G81891 hypothetical prote  
35 29 14.5 35 2 S21191 cytochrome c550 -  
36 28.5 14.2 21 2 PC4381 dehydrin 7.5K poly  
37 28.5 14.2 23 2 B30988 hypothetical prote  
38 28.5 14.2 26 2 PQ0105 alkaline serine pr  
39 28.5 14.2 27 2 F70540 hypothetical prote  
40 28.5 14.2 32 4 A49016 gag-junD mutant fu  
41 28 14.0 10 2 XGHUE erythrocyte membra  
42 28 14.0 15 2 PH1582 ig H chain V-D-J r  
43 28 14.0 23 2 T44539 hypothetical prote  
44 28 14.0 23 2 S26243 reed protein - Str  
45 28 14.0 23 2 A32473 histidine-rich pro

## ALIGNMENTS

## RESULT 1

A41683  
hyaluronate receptor CD44 (version 2) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: A41683  
R:Shtivelman, E.; Bishop, J.M.  
Mol. Cell. Biol. 11, 5446-5453, 1991  
A:Title: Expression of CD44 is repressed in neuroblastoma cells.  
A:Reference number: A41683; MUID:92017823  
A:Accession: A41683  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-29 <SHT>  
A:Cross-references: GB:M69215  
C:Keywords: cell adhesion

Query Match 31.5% Score 63; DB 2; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Query 5 DQDTFHPSGGS 15

Db 19 DQDTFHPSGGS 29  
|||||

## RESULT 2

C31514  
hemopexin - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 18-Jun-1993  
C:Accession: C31514  
R:Wellner, D.; Cheng, K.C.; Muller-Eberhard, U.  
Biochem. Biophys. Res. Commun. 155, 622-625, 1988  
A:Title: N-terminal amino acid sequences of the hemopexins from chicken, rat and rabb  
A:Reference number: A90148; MUID:88339942  
A:Accession: C31514  
A:Molecule type: protein  
A:Residues: 1-34 <WE3>

Query Match 21.0% Score 42; DB 2; Length 34;  
Best Local Similarity 35.0%; Pred. No. 52;  
Matches 7; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

## Query 16 HTHTESEDHSHGSGEGGAN 35

Db 6 HKHTPGDEHPHGAEPDCXD 25  
| | | | | | | | | |

## RESULT 3

PQ0548  
 capsid protein VP26 - human herpesvirus 1 (fragment)  
 C:Species: human herpesvirus 1  
 C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
 C:Accession: PQ0548  
 R:Davidson, M.D.; Rixon, F.J.; Davison, A.J.  
 J. Gen. Virol. 73, 2709-2713, 1992  
 A:Title: Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes  
 A:Reference number: PQ0544; MUID:93019027  
 A:Accession: PQ0548  
 A:Molecule type: protein  
 A:Residues: 1-19 <DAV>  
 A:Experimental source: strain 17  
 C:Genetics:  
 A:Gene: UL35  
 C:Keywords: capsid protein

Query Match 19.0%; Score 38; DB 2; Length 19;  
 Best Local Similarity 58.3%; Pred. No. 92;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATRDQTFHPG 13  
 || :|| || |  
 Db 6 ATNODNPHPG 17

## RESULT 4

PQ0263  
 dnaK-type molecular chaperone BLP-3 - common tobacco (fragment)  
 N:Alternate names: luminal binding protein BLP-3  
 C:Species: Nicotiana tabacum (common tobacco)  
 C>Date: 30-Jun-1992 #sequence\_revision 10-Mar-1994 #text\_change 13-Mar-1998  
 C:Accession: PQ0263  
 R:Denecke, J.; Goldman, M.H.S.; Demolder, J.; Seurinck, J.; Botterman, J.  
 Plant Cell 3, 1025-1035, 1991  
 A:Title: The tobacco luminal binding protein is encoded by a multigene family.  
 A:Reference number: JQ1360; MUID:92361242  
 A:Accession: PQ0263  
 A:Molecule type: mRNA  
 A:Residues: 1-28 <DEN>

A:Note: translation of the nucleotide sequence is not complete  
 C:Comment: This protein plays a role both in protein import into the lumen of the endoplasmic reticulum and in protein folding and assembling/disassembling of protein complex  
 C:Genetics:  
 A:Gene: BLP-3  
 C:Function:  
 A:Description: involved in protein folding and assembling/disassembling of protein complex  
 C:Superfamily: heat shock protein 70  
 C:Keywords: ATP; molecular chaperone  
 F:25-28/Region: endoplasmic reticulum retention signal

Query Match 17.5%; Score 35; DB 2; Length 28;  
 Best Local Similarity 57.1%; Pred. No. 3.3e+02;  
 Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 12 SGSSHTTHEEDGH 25  
 ||||| :|||  
 Db 14 SGSSSS--EEDGH 25

## RESULT 5

T04371  
 thaumatin-like protein - barley (fragment)  
 C:Species: Hordeum vulgare (barley)  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T04371  
 R:Skadsen, R.W.; Herbst, J.M.  
 submitted to the EMBL Data Library, July 1997  
 A:Reference number: Z15316  
 A:Accession: T04371  
 A:Status: Preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA

A:Residues: 1-26 <SKA>  
 A:Cross-references: EMBL:AF016328; NID:g2454603; PIDN:AAB71681.1; PID:g2454604  
 A:Experimental source: cv. Morex  
 C:Genetics:  
 A:Gene: perm2

Query Match 17.2%; Score 34.5; DB 2; Length 26;  
 Best Local Similarity 43.5%; Pred. No. 3.5e+02;  
 Matches 10; Conservative 1; Mismatches 7; Indels 5; Gaps 2;

QY 11 PSGSHTTHESEDEGHSGSQEGG 33  
 ||| :||| || |  
 Db 3 PGGG----HSSPPPGHG-GRQDLG 20

## RESULT 6

I29501  
 fibrinopeptide A - kangaroo  
 C:Species: Macropus sp. (kangaroo)  
 C>Date: 21-Nov-1987 #sequence\_revision 08-Jun-1990 #text\_change 18-Aug-2000  
 C:Accession: I29501  
 R:Blombaeck, B.; Blombaeck, M.; Hann, C.  
 unpublished results, cited by Blombaeck, B., and Blombaeck, M., in Chemotaxonomy and  
 A:Reference number: A29501  
 A:Accession: I29501  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <BLO>

C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology  
 Query Match 16.5%; Score 33; DB 2; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 TRDQDTFHPGG 14  
 ||| :|| |  
 Db 1 TKDSGTFAEGG 12

## RESULT 7

Kex2/subtilisin-like proprotein convertase PC4-C - mouse (fragment)  
 F45357  
 C:Species: Mus musculus (house mouse)  
 C>Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
 C:Accession: F45357  
 R:Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.  
 Mol. Endocrinol. 6, 1559-1570, 1992  
 A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ  
 A:Reference number: A45357; MUID:93078790  
 A:Accession: F45357  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-24 <SET>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118888)

Query Match 16.5%; Score 33; DB 2; Length 24;  
 Best Local Similarity 60.0%; Pred. No. 5e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 13 GGSHTTHESE 22  
 ||| :|| |  
 Db 15 GGSTATSSQ 24

## RESULT 8

C45357  
 Kex2/subtilisin-like proprotein convertase PC4-C - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
 C:Accession: C45357

R;Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.  
 Mol. Endocrinol. 6, 1559-1570, 1992  
 A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell protein  
 A:Reference number: A45357; MUID:93078790  
 A:Accession: C45357  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-24 <SEI>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118879)

Query Match 16.5%; Score 33; DB 2; Length 24;  
 Best Local Similarity 60.0%; Pred. No. 5e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GGSHTTHESE 22  
 ||| |||  
 Db 15 GGSTATSHSQ 24

RESULT 9  
 PC2342  
 second envelope glycoprotein, gp70, hypervariable region 1 - hepatitis C virus (isolate)  
 C:Species: hepatitis C virus  
 C:Date: 21-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 08-Oct-1999  
 C:Accession: PC2342  
 R;Kato, N.; Nakazawa, T.; Mizutani, T.; Shimotohno, K.  
 Biochem. Biophys. Res. Commun. 206, 863-869, 1995  
 A:Title: Susceptibility of human T-lymphotropic virus type I infected cell line MT-2 to  
 A:Reference number: PC2334; MUID:95134269  
 A:Accession: PC2342  
 A:Molecule type: genomic RNA  
 A:Residues: 1-27 <KAT>  
 A:Cross-references: DDBJ:D43654; NID:G882093; PIDN:BAA07769.1; PID:di008354; PID:g882094  
 A:Experimental source: isolate D-1, inoculum  
 C:Keywords: glycoprotein

Query Match 16.5%; Score 33; DB 2; Length 27;  
 Best Local Similarity 46.2%; Pred. No. 5.6e+02;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 17 TTSEEDGSHSGS 29  
 ||| | :||  
 Db 1 TTHVSGGAQAHA 13

RESULT 10  
 152232  
 tau protein - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999  
 C:Accession: 152232  
 R;Mori, H.; Hamada, Y.; Kawaguchi, M.; Honda, T.; Kondo, J.; Ihara, Y.  
 Biochem. Biophys. Res. Commun. 159, 1221-1226, 1989  
 A:Title: A distinct form of tau is selectively incorporated into Alzheimer's paired helix  
 A:Reference number: 152232; MUID:89193714  
 A:Accession: 152232  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-31 <RES>  
 A:Cross-references: GB:M25298; NID:G602470; PIDN:AAA57264.1; PID:g602471  
 C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology

Query Match 16.5%; Score 33; DB 2; Length 31;  
 Best Local Similarity 40.0%; Pred. No. 6.5e+02;  
 Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CATRDQDTFHPGSGS 15  
 |::| |::|  
 Db 17 CGSKDNKIKHVPGGGS 31

## RESULT 11

E45357  
 Kex2/subtilisin-like proprotein convertase PC4-B - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
 C:Accession: E45357  
 R;Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.  
 Mol. Endocrinol. 6, 1559-1570, 1992  
 A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ  
 A:Reference number: A45357; MUID:93078790  
 A:Accession: E45357  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-34 <SEI>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118887)

Query Match 16.5%; Score 33; DB 2; Length 34;  
 Best Local Similarity 60.0%; Pred. No. 7.1e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GGSHTTHESE 22  
 ||| |||  
 Db 25 GGSTATSHSQ 34

RESULT 12  
 B45357  
 Kex2/subtilisin-like proprotein convertase PC4-B - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
 C:Accession: B45357  
 R;Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.  
 Mol. Endocrinol. 6, 1559-1570, 1992  
 A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ  
 A:Reference number: A45357; MUID:93078790  
 A:Accession: B45357  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-34 <SEI>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118877)

Query Match 16.5%; Score 33; DB 2; Length 34;  
 Best Local Similarity 60.0%; Pred. No. 7.1e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GGSHTTHESE 22  
 ||| |||  
 Db 25 GGSTATSHSQ 34

RESULT 13  
 B32473  
 histidine-rich protein C, peptide p-5 - liver fluke (fragment)  
 C:Species: Fasciola hepatica (liver fluke)  
 C:Date: 25-Sep-1989 #sequence\_revision 03-May-1994 #text\_change 26-May-2000  
 C:Accession: B32473  
 R;Waite, J.H.; Rice-Ficht, A.C.  
 Biochemistry 28, 6104-6110, 1989  
 A:Title: A histidine-rich protein from the vitellaria of the liver fluke Fasciola hep  
 A:Reference number: A32473; MUID:89375343  
 A:Accession: B32473  
 A:Molecule type: protein  
 A:Residues: 1-18 <WAI>  
 A:Note: 18-Gly and 18-His were also found  
 C:Superfamily: period clock protein; EGF homology  
 C:Keywords: egg yolk  
 F,1,5/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental

Query Match 16.0%; Score 32; DB 2; Length 18;

Best Local Similarity 42.9%; Pred. No. 5e+02;  
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 19 HESEDGSHGSGEG 32  
| | | | |  
DB 3 HGYGHHGHGHG 16

## RESULT 14

B82151  
hypothetical protein Vcl830 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: B82151  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: B82151  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-31 <HEI>  
A:Cross-references: GB:AE004259; GB:AE003852; NID:g9656353; PIDN:AAF94978.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: Vcl830  
A:Map position: 1

Query Match 16.0%; Score 32; DB 2; Length 31;

Best Local Similarity 50.0%; Pred. No. 8.6e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 21 SEDGHSH--GSQEG 32  
: | | : | | |  
DB 3 AELGHTHKKKREG 16

## RESULT 15

JP0101  
fibrinogen alpha chain - duck (fragment)  
N:Contains: fibrinopeptide A  
C:Species: Anas platyrhynchos (domestic duck)  
C:Date: 30-Jun-1987 #sequence\_revision 28-Dec-1987 #text\_change 26-Jan-1996  
C:Accession: JP0101  
R:Min, Y.; Ping, Z.; Yaoshi, Z.  
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 28, 31-35, 1985  
A:Title: Purification and primary structures of duck fibrinopeptides A and B.  
A:Reference number: A94238  
A:Accession: JP0101  
A:Molecule type: protein  
A:Residues: 1-15 <MIN>  
C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology  
C:Keywords: blood coagulation; plasma; pyroglutamic acid  
F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 15.5%; Score 31; DB 2; Length 15;

Best Local Similarity 50.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 22 EDGSHGSGEGG 33  
: | | : | | |  
DB 1 QDGKSSFGKEG 12

Search completed: February 21, 2002, 16:53:30  
Job time: 159 sec



Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	42	21.0	34	1	HEMO_CHICK	P20057	gallus galli
2	41	15.5	15	1	FIBA_ANAPL	P12801	anas platyr
3	31	15.5	19	1	FTBB_HORSE	P14471	equus cabal
4	31	15.5	23	1	FMK7_PSEAE	P53391	pseudomonas
5	30.5	15.2	32	1	YL55_CAEEL	P34435	caenorhabdi
6	29	14.5	11	1	COXA_CANFA	P99501	canis famill
7	29	14.5	33	1	HF40_MAZE1	P82865	zea mays (m
8	29	14.5	35	1	C550_BACHA	P80091	bacillus ha
9	28	14.0	10	1	GLEM_HUMAN	P02728	homo sapien
10	28	14.0	24	1	C0XC_THUOB	P80973	thunnus obe
11	28	14.0	32	1	UC09_MAZE1	P80615	zea mays (m
12	27.5	13.8	29	1	RL15_HALCU	P05971	halobacteri
13	27	13.5	13	1	CXA2_CONGE	P01520	conus geogr
14	27	13.5	15	1	FIBA_SYNGA	P14463	syncerus ca
15	27	13.5	22	1	DP20_CAEBR	P51558	caenorhabdi
16	27	13.5	28	1	LECA_TRIHO	P36230	iris hollan
17	27	13.5	33	1	PEN3_ADECU	P35987	canine aden
18	26	13.0	20	1	FRE3_LITIN	P56429	litorea inf
19	25	12.5	16	1	R1PK_TRIKI	P16093	trichosanthe
20	25	12.5	18	1	CXA1_CONER	P50982	conus ermin
21	25	12.5	19	1	FRHG_METBA	P80491	methanosarc
22	25	12.5	20	1	C0XM_THUOB	P80981	thunnus obe
23	25	12.5	20	1	MCRG_METTE	P22950	methanosarc
24	25	12.5	20	1	PORC_METTM	P80902	methanobact
25	25	12.5	23	1	CHTB_PEA	P21227	pisum sativ
26	25	12.5	25	1	DNAB_MYCCA	P71500	mycoplasma
27	25	12.5	31	1	EFTU_STRLU	P52390	streptomyce
28	25	12.5	33	1	THIO_CLOST	P81109	clostridium
29	25	12.5	34	1	VPU_HV1W2	P08808	human immun
30	24	12.0	10	1	AMFN_HELAM	P81731	helicoverpa
31	24	12.0	14	1	FIBA_HORSE	P14452	equus cabal
32	24	12.0	20	1	COGA_CHIOP	P34156	chionocete
33	24	12.0	21	1	PEDB_HYDAT	P80577	hydra atten

DE Fibropneumoniae (Domestic duck).  
OS Anas platyrhynchos (Domestic duck).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.  
OX NCBI\_TaxID=8839;  
RN [1]

```

RP SEQUENCE=85168193; PubMed-3983613;
RA Min Y., Ping Z., Yaoshi Z.;
RT "Purification and primary structures of duck fibrinopeptides A and B.";
RL Sci. Sin., B, Chem. Biol. Agric. Med. Earth Sci. 28:31-35(1985).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC PIR; JP0101; JP0101.
KW Blood coagulation; Plasma.
DR MOD_RES 1 15 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1580 MW; D78A51FF88B40373 CRC64;

Query Match 15.5%; Score 31; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 22 EDGSHSGSQEGG 33
:| | :| |
Db 1 QDGKSSFOKEGG 12

RESULT 3
FIBB_HORSE
ID FIBB_HORSE STANDARD; PRT; 19 AA.
AC P14471;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 01-JAN-1990 (Rel. 13, Last annotation update)
DE FIBRINOPEPTIDE B.
OS Equus caballus (Horse), and Equus asinus (Donkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796, 9793;
RN [1]
RP SEQUENCE=Horse;
RC Blomback B., Blomback M., Grondahl N.J.; "
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN [2]
RP SEQUENCE.
RC SPECIES=E.asinus;
RA Blomback B., Blomback M., Grondahl N.J., Holmberg E.;
RT "Structure of fibrinopeptides-its relation to enzyme specificity and phylogeny and classification of species.";
RL Ark. Kemi 25:411-428(1966).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
KW PROSITE; PS00514; FIBRIN_AG_C DOMAIN; PARTIAL.
DR Blood coagulation; Plasma; Sulfation.
FT MOD_RES 3 3
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2296 MW; 921A2B02D5F6891D CRC64;

```

```

Query Match 15.5%; Score 31; DB 1; Length 19;
Best Local Similarity 62.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 19 HESEDGHS 26
:| | | | :
Db 5 HEEDGRT 12

RESULT 4
FMK7_PSEAE
ID FMK7_PSEAE STANDARD; PRT; 23 AA.
AC Q53391;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE FIMBRIAL PROTEIN (PILIN) (STRAIN KB7) (FRAGMENT).
GN PIL.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KB7;
RX MEDLINE=94103636; PubMed=7903973;
RA IPATSG;
RT "A multicenter comparison of methods for typing strains of Pseudomonas aeruginosa predominant from patients with cystic fibrosis. The international Pseudomonas aeruginosa Typing Study Group.";
RL J. Infect. Dis. 169:134-142(1994).
RN [2]
RP STRUCTURE BY NMR OF 7-23.
RC STRAIN=KB7;
RX MEDLINE=96110702; PubMed=8845350;
RA Campbell A.P., McInnes C., Hodges R.S., Sykes B.D.;
RT "Comparison of NMR solution structures of the receptor binding domains of Pseudomonas aeruginosa pili strains PAO, KB7, and PAK: implications for receptor binding and synthetic vaccine design.";
RL Biochemistry 34:16255-16268(1995).
CC -!- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4 NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.
CC -!- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S67809; CAB32861.1;
DR PDB; 1KB7; 29-JAN-96.
DR PDB; 1KB8; 29-JAN-96.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyltn.
DR Pfam; PF00114; pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; PARTIAL.
KW Fimbria; 3D-structure.
FT NON_TER 1 1
FT DISULFID 8 21
SQ SEQUENCE 23 AA; 2415 MW; D0BE77514AF041CF CRC64;

Query Match 15.5%; Score 31; DB 1; Length 23;
Best Local Similarity 46.2%; Pred. No. 4e+02;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CATRDQDTFHPG 13

```



```

OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE, AND STRUCTURE BY NMR.
RC STRAIN-SP. NOV./ ATCC 49067;
RX MEDLINE-92201200; PubMed-1312933;
RA Saraiva L.M., Denariáz G., Liu M.-Y., Payne W.J., le Gall J.,
RA Moura I.;
RT "NMR and EPR studies on a monoheme cytochrome c550 isolated from
RT Bacillus halodentrificans";
RL Eur. J. Biochem. 204:1131-1139(1992).
CC -!- FUNCTION: MONOHEME CYTOCHROME WHICH FUNCTIONS AS AN ELECTRON
CC CARRIER IN THE REDUCTION OF NITRITE BY MEMBRANE VESICLES.
CC -!- PTM: BINDS ONE HEME GROUP PER MOLECULE.
DR PIR: S21191; S21191.
DR InterPro: IPR000345; CytC_heme_bind.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
KW Electron transport; Heme.
FT BINDING 17 17 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 20 20 HEME (COVALENT) (BY SIMILARITY).
FT METAL 21 21 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3352 MW; C3B5D6D8F23004EB CRC64;

Query Match 14.5%; Score 29; DB 1; Length 35;
Best Local Similarity 30.4%; Pred. No. 1.1e+03;
Matches 7; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 12 SGGSHTHSEDSHGSHGSEGA 34
| : || ||: ||
Db 6 SAAAEVFPESCASGADLSGA 28

RESULT 9
GLEM_HUMAN
ID GLEM_HUMAN STANDARD; PRT; 10 AA.
AC P02728;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ERYTHROCYTE MEMBRANE GLYCOPOLYPEPTIDE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE-72034940; PubMed=5286858;
RA Weiss J.B., Lote C.J., Bobinski H.;
RT "New low molecular weight glycopeptide containing triglucoylcysteine
RT in human erythrocyte membrane.";
RL Nature New Biol. 234:25-26(1971).
CC -!- PTM: S-LINKED GLYCAN CONSISTS OF GLC-GLC-GLC TRISACCHARIDE.
CC -!- MISCELLANEOUS: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS
CC PEPTIDE IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN
CC ATTRIBUTED.
DR PIR: A03187; XGHUE.
KW Glycoprotein; Erythrocyte.
FT CARBOHYD 1 1 S-LINKED (GLC...);
SQ SEQUENCE 10 AA; 1049 MW; 239BFEEA1F5B1E8 CRC64;

Query Match 14.0%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 23 DGHSH 27
| : ||||
Db 2 EGHSH 6

```

```

RESULT 10
COXC_THUOB
ID COXC_THUOB STANDARD; PRT; 24 AA.
AC P80973;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-2 (EC 1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE-97454291; PubMed-9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2903 MW; 20998FB91F22B43B CRC64;

Query Match 14.0%; Score 28; DB 1; Length 24;
Best Local Similarity 83.3%; Pred. No. 9.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 SHGSQE 31
||| ||
Db 1 SHGKQE 6

RESULT 11
UC09_MAIZE
ID UC09_MAIZE STANDARD; PRT; 32 AA.
AC P80615;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 206)
DE (FRAGMENTS).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.4, ITS MW IS: 16.9 KDA.
CC -!- SIMILARITY: TO THE TOMATO ABCISIC STRESS RIPENING PROTEINS.
CC -!- CAUTION: THE ORDER OF THE PEPTIDES SHOWN IS UNCERTAIN.
DR Maize-2DPAGE; P80615; COLEOPTILE.
DR MaizeDB; 123936; -.
FT NON_TER 1 1
FT NON_CONS 17 18

```

FT NON\_CONS 24 25  
 FT NON\_TER 32 32  
 SQ SEQUENCE 32 AA; 3249 MW; 8841DBE253F211D6 CRC64;

Query Match 14.0%; Score 28; DB 1; Length 32;  
 Best Local Similarity 35.7%; Pred. No. 1.3e+03;  
 Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 14 GSHTHSESDGSHG 27  
 Db 14 GGYVXXLGEAGHHH 27

## RESULT 12

RL15\_HALCU STANDARD; PRT; 29 AA.  
 ID RL15\_HALCU  
 AC P05971;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L15P (HL16) (FRAGMENT).  
 GN RPL15P.  
 OS Halobacterium cutirubrum.  
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
 CC Halobacterium.  
 OX NCBI\_TaxID=2242;  
 RN [1]  
 RP SEQUENCE.

RA MEDLINE=84282108; PubMed=6467081;  
 RA Matheson A.T., Yaguchi M., Christensen P., Rollin C.F., Hasnain S.;  
 RT "Purification, properties, and N-terminal amino acid sequence of  
 certain 50S ribosomal subunit proteins from the archaeobacterium  
 Halobacterium cutirubrum.";  
 RT Can. J. Biochem. Cell Biol. 62:426-433(1984).  
 CC -!- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.  
 DR PIR: S08555; S08555.  
 DR InterPro: IPR001196; Ribosomal\_L15.  
 DR PROSITE: PS00475; RIBOSOMAL\_L15; PARTIAL.  
 KW Ribosomal protein.  
 FT NON\_TER 29  
 SQ SEQUENCE 29 AA; 3170 MW; 73D13D831B4BE790 CRC64;

Query Match 13.8%; Score 27.5; DB 1; Length 29;  
 Best Local Similarity 37.5%; Pred. No. 1.4e+03;  
 Matches 6; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 13 GGSHTHSESDGSHG 28  
 Db 15 GGTHKNSRGA-GNRHG 29

## RESULT 13

CXA2\_CONGE STANDARD; PRT; 13 AA.  
 ID CXA2\_CONGE  
 AC P01520;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1988 (Rel. 09, Last annotation update)  
 DE ALPHA-CONOTOXIN GII.  
 OS Conus geographus (Geography cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 CC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=6491;  
 RN [1]  
 RP SEQUENCE.

RA MEDLINE=81191854; PubMed=7014556;  
 RA Gray W.R., Luque A., Olivera B.M., Barrett J., Cruz L.J.;  
 RT "Peptide toxins from Conus geographus venom.";  
 RL J. Biol. Chem. 256:4734-4740(1981).  
 RN [2]  
 RP DISULFIDE BONDS.

RX MEDLINE=84032400; PubMed=6630187;  
 RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;  
 RT "Conotoxin M1. Disulfide bonding and conformational states.";  
 RL J. Biol. Chem. 258:12247-12251(1983).  
 RN [3]  
 RP REVIEW.

RX MEDLINE=89024586; PubMed=3052286;  
 RA Gray W.R., Olivera B.M., Cruz L.J.;  
 RT "Peptide toxins from venomous Conus snails.";  
 RL Annu. Rev. Biochem. 57:665-700(1988).  
 CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
 CC INHIBIT THEM.

DR PIR: A01783; NTKN2G.  
 DR HSSP: P01519; INOT.  
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
 KW Venom.

FT DISULFID 2 7  
 FT DISULFID 3 13  
 FT MOD\_RES 13 13  
 SQ SEQUENCE 13 AA; 1422 MW; DEEE831C39297EBD CRC64;

Query Match 13.5%; Score 27; DB 1; Length 13;  
 Best Local Similarity 57.1%; Pred. No. 7e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 HPSGGSH 16  
 Db 4 HPACGKH 10

## RESULT 14

FIBA\_SYNCA STANDARD; PRT; 15 AA.  
 ID FIBA\_SYNCA  
 AC P14463;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-JAN-1990 (Rel. 13, Last annotation update)  
 DE FIBRINOPEPTIDE A.  
 OS Syncerus caffer (Cape buffalo).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Syncerus.  
 OX NCBI\_TaxID=9970;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=67209145; PubMed=6033721;  
 RA Doolittle R.F., Schubert D., Schwartz S.A.;  
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.  
 RT Dromedary camel, mule deer, and cape buffalo.";  
 RL Arch. Biochem. Biophys. 118:456-467(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.

CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 KW Blood coagulation; Plasma.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1480 MW; 4E998EA5F0B41CC6 CRC64;

Query Match 13.5%; Score 27; DB 1; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 8.1e+02;  
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 22 EDGSHGSHQEGG 33  
 Db 1 EDGSGEFLAEGG 12

```

RESULT 15
DP20_CAEBR
ID DP20_CAEBR STANDARD; PRT; 22 AA.
AC P51S58;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE DPY-20 PROTEIN (FRAGMENT).
GN DPY-20.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95287857; PubMed=7770042;
RA Clark D.V., Suleman D.S., Beckenbach K.A., Gilchrist E.J.,
RA Baillie D.L.;
RT "Molecular cloning and characterization of the dpy-20 gene of
RT Caenorhabditis elegans."
RL Mol. Gen. Genet. 247:367-378(1995).
CC -!- FUNCTION: INVOLVED IN CUTICLE FUNCTION AND IS ESSENTIAL FOR
CC NORMAL MORPHOLOGICAL DEVELOPMENT (BY SIMILARITY).
KW Cuticle.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2529 MW; A11FB717AD367F69 CRC64;

Query Match 13.5%; Score 27; DB 1; Length 22;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 23 DGHSHGS 29
Db 2 DGHSNST 8

```

Search completed: February 21, 2002, 16:56:16  
 Job time: 184 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:52:56 ; Search time 21.89 Seconds  
(without alignments)  
233.875 Million cell updates/sec

Title: US-08-753-851-8  
Perfect score: 200  
Sequence: 1 CATRDQDTFHPGSGSHTHESEDSHGSGOEGGAN 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 16942

Minimum DB seq length: 0  
Maximum DB seq length: 35

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	23.0	27	12	037250 hepatitis c
2	39	19.5	26	12	011897 hepatitis g
3	39	19.5	27	12	056472 hepatitis c
4	39	19.5	27	12	090900 hepatitis c
5	39	19.5	27	12	090907 hepatitis c
6	39	19.5	27	12	090905 hepatitis c
7	38	19.0	17	12	078324 human immun
8	38	19.0	17	12	078378 human immun
9	38	19.0	21	12	09YOS8 porcine cir
10	38	19.0	21	12	056126 porcine cir
11	38	19.0	26	12	09QRT5 hepatitis c
12	38	19.0	27	12	0912X5 hepatitis c
13	38	19.0	28	9	Q37931 bacterioph
14	37	18.5	17	12	078381 human immun
15	36	18.0	17	12	078345 human immun
16	36	18.0	21	12	093191 porcine cir
17	36	18.0	27	12	056521 hepatitis c
18	36	18.0	27	12	09J5V4 hepatitis c
19	36	18.0	30	10	Q96240 arabidopsis

20	35	17.5	17	12	078323 human immun
21	35	17.5	17	12	078379 human immun
22	35	17.5	17	12	078327 human immun
23	35	17.5	17	12	078380 human immun
24	35	17.5	27	12	09J5W1 hepatitis c
25	35	17.5	27	12	09J5V7 hepatitis c
26	34	17.2	26	10	022463 hordeum vul
27	34	17.0	21	12	078511 human immun
28	34	17.0	25	12	011890 hepatitis g
29	34	17.0	26	12	011894 hepatitis c
30	34	17.0	27	12	09J5W0 hepatitis c
31	34	17.0	27	12	09IPH8 hepatitis c
32	34	17.0	27	12	09IPH5 hepatitis c
33	34	17.0	27	12	09ILE5 hepatitis c
34	34	17.0	27	12	09ILD9 hepatitis c
35	34	17.0	27	12	09ILD6 hepatitis c
36	34	17.0	27	12	09ILD3 hepatitis c
37	34	17.0	27	12	09ILD0 hepatitis c
38	34	17.0	27	12	09ILC8 hepatitis c
39	34	17.0	27	12	037263 hepatitis c
40	34	17.0	29	4	043807 homo sapien
41	33.5	16.8	27	12	056479 hepatitis c
42	33.5	16.8	27	12	056526 hepatitis c
43	33	16.5	16	12	078510 human immun
44	33	16.5	20	12	078486 human immun
45	33	16.5	20	12	078508 human immun

#### ALIGNMENTS

RESULT 1  
O37250 PRELIMINARY; PRT; 27 AA.  
AC O37250;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
GN E2.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=15;  
RX MEDLINE=98105815; PubMed=9445070;  
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;  
RT "Genetic diversity and tissue compartmentalization of the hepatitis C  
RT virus genome in blood mononuclear cells, liver, and serum from chronic  
RT hepatitis C patients.";  
RL J. Virol. 72:1640-1646(1998).  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NS1 IS  
CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
CC PROTEIN C: THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
CC FAMILY.  
CC EMBL; AF018359; AAC03646.1; -.  
CC InterPro; IPR002531; HCV\_NSL1.  
CC Pfam; PF01560; HCV\_NSL1; 1.  
CC Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2820 MW; 47C59C48D18F3DC7 CRC64;





OS Hepatitis C virus.  
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20091320; PubMed=10623727;  
 RA Sandres K., Dubois M., Pasquier C., Payen J.L., Alric L., Duffaut M.,  
 RA Vinet J.P., Pascal J.P., Puel J., Izopet J.;  
 RT "Genetic heterogeneity of hypervariable region 1 of the hepatitis C  
 RT virus (HCV) genome and sensitivity of HCV to alpha interferon  
 RT therapy.";  
 RL J. Virol. 74:661-668(2000).  
 CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
 CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
 CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
 CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
 CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
 CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
 CC FAMILY.  
 CC EMBL; AF166573; AAD52231.1; .  
 DR InterPro: IPR002531; HCV\_NSI.1.  
 DR Pfam: PF01560; HCV\_NSI.1.  
 DR KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 2934 MW; CB80D3F44ECBEAC4 CRC64;  
 Query Match 19.5%; Score 39; DB 12; Length 27;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
 Matches 11; Conservative 1; Mismatches 8; Indels 2; Gaps 1;  
 RESULT 6  
 QY 12 SGG--SHTTHESEGDHSGSQE 31  
 DB 5 SGGTASHTTTRRLTSLFSFGSQ 26  
 AC Q909B5 PRELIMINARY; PRT; 27 AA.  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
 DE (FRAGMENT).  
 OS Hepatitis C virus.  
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20091320; PubMed=10623727;  
 RA Sandres K., Dubois M., Pasquier C., Payen J.L., Alric L., Duffaut M.,  
 RA Vinet J.P., Pascal J.P., Puel J., Izopet J.;  
 RT "Genetic heterogeneity of hypervariable region 1 of the hepatitis C  
 RT virus (HCV) genome and sensitivity of HCV to alpha interferon  
 RT therapy.";  
 RL J. Virol. 74:661-668(2000).  
 CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
 CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
 CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
 CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
 CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
 CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
 CC FAMILY.  
 CC EMBL; AF166585; AAD52243.1; .

DR InterPro: IPR002531; HCV\_NSI.1.  
 DR Pfam: PF01560; HCV\_NSI.1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 2904 MW; CB80C3F54ECBEAC4 CRC64;  
 Query Match 19.5%; Score 39; DB 12; Length 27;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
 Matches 11; Conservative 1; Mismatches 8; Indels 2; Gaps 1;  
 RESULT 7  
 QY 12 SGG--SHTTHESEGDHSGSQE 31  
 DB 5 SGGTASHTTTRRLTSLFSFGSQ 26  
 AC Q78324 PRELIMINARY; PRT; 17 AA.  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE VIRAL SAMPLE FLIPR5B (FLORIDA PATIENT A), PARTIAL ENV CDS, V5 REGION  
 DE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang L.Q., Leigh-Brown A.J.;  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92271245; PubMed=1589796;  
 RA Ou C.Y., Ciesielski C.A., Myers G., Banda C.I., Luo C.C.,  
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,  
 RA Economidou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,  
 RA Curran J.W., Jaffe H.W.;  
 RT "Molecular epidemiology of HIV transmission in a dental practice.";  
 RL Science 256:1165-1171(1992).  
 DR EMBL; M92110; AAA44466.1; .  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1651 MW; 3473A0BB802CA370 CRC64;  
 Query Match 19.0%; Score 38; DB 12; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 RESULT 8  
 QY 3 TRDQDTFHPGG 14  
 DB 5 TKGSETFPGGG 16  
 AC Q78378 PRELIMINARY; PRT; 17 AA.  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE VIRAL SAMPLE FLIPR5A (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION  
 DE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang L.Q., Leigh-Brown A.J.;

```

RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Banea C.I., Luo C.C.,
RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
RA Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92123; AAA44493.1; -
FT NON_TER 1
FT 17
SQ SEQUENCE 17 AA; 1723 MW; 34757935D12CA370 CRC64;

Query Match 19.0%; Score 38; DB 12; Length 17;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 TRDQTFHPGG 14
Db 5 TNETETFRPGG 16
| : || | |
| : || | |

RESULT 9
QYQ58 PRELIMINARY; PRT; 21 AA.
AC QYQ58;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE ORF-8.
OS Porcine circovirus type 2-D.
OC Viruses; ssDNA viruses; Circoviridae; Circovirus.
OX NCBI_TaxID=86385;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamel A.L., Nayar G.P.S.;
RT "Genetic characterization of four novel type-2 Porcine circoviruses.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117753; AAD12312.1; -
SQ SEQUENCE 21 AA; 2301 MW; BD26E338284C1013 CRC64;

Query Match 19.0%; Score 38; DB 12; Length 21;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 HPSGGSHHTHE 20
Db 11 HPTAASHKSHQ 21
| : || : |
| : || : |

RESULT 10
QYQ58 PRELIMINARY; PRT; 21 AA.
AC QYQ58;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE STRAIN PMWS PCV, COMPLETE GENOME (ORF-8).
OS porcine circovirus,
OS bovine circovirus,
OS porcine circovirus type 2-E, and
OS porcine circovirus type 2-B.
OC Viruses; ssDNA viruses; Circoviridae; Circovirus.
OX NCBI_TaxID=46221, 85542, 85544, 85709;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=porcine circovirus; STRAIN=PMWS PCV;
RX MEDLINE=98241172; PubMed=9573301;
RA Hamel A.L., Lin L.L., Nayar G.P.S.;

```

```

RT "Nucleotide sequence of porcine circovirus associated with postweaning
multisystemic wasting syndrome in pigs.";
J. Virol. 72:5262-5267(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=porcine circovirus; STRAIN=PMWS PCV;
RA Hamel A.L., Lin L.L., Nayar G.P.S.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=porcine circovirus; STRAIN=PORCINE CIRCOVIRUS TYPE II;
RX MEDLINE=98418498; PubMed=9747726;
RA Meehan B.M., McNeilly F.M., Todd D., Kennedy S., Jewhurst V.,
RA Ellis J.A., Hassard L.E., Clark E.G., Haines D.M., Allan G.M.;
RT "Characterization of novel circovirus DNAs associated with wasting
syndromes in pigs.";
J. Gen. Virol. 79:2171-2199(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=porcine circovirus; STRAIN=PORCINE CIRCOVIRUS TYPE II;
RA Meehan B.M.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=bovine circovirus;
RA Hamel A.L., Nayar G.P.S.;
RT "Nucleotide sequence of a circovirus detected in cattle with various
clinical syndromes.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=porcine circovirus type 2-E;
RA Hamel A.L., Nayar G.P.S.;
RT "Nucleotide sequence of four different isolates of circovirus detected
in pigs with various clinical syndromes.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=porcine circovirus type 2-B;
RA Hamel A.L., Nayar G.P.S.;
RT "Nucleotide sequence of four different isolates of porcine circovirus
detected in pigs with various clinical syndromes.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF027217; AAC59469.1; -
DR EMBL; AF055392; AAC35314.1; -
DR EMBL; AF055391; AAC35303.1; -
DR EMBL; AF109397; AAD11935.1; -
DR EMBL; AF109399; AAD03078.1; -
DR EMBL; AF112862; AAD03094.1; -
SQ SEQUENCE 21 AA; 2313 MW; BD26F6E2184C1013 CRC64;

Query Match 19.0%; Score 38; DB 12; Length 21;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 HPSGGSHHTHE 20
Db 11 HPTAASHKSHQ 21
| : || : |
| : || : |

RESULT 11
QYQ58 PRELIMINARY; PRT; 26 AA.
AC QYQ58;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]
DE (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

```

```

OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G-SE;
RA Yeh C.-Y.;
RT "Replication of hepatitis C virus in the ascitic mononuclear cells and
RL development of distinct quasiespecies in the ascitic fluid.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NS1 IS
CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY
CC SIMILARITY).
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1
CC FAMILY.
CC EMBL; AF109756; AAD51585.1; -.
DR InterPro: IPR002531; HCV_NSI.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 26
FT NON_TER 27
SQ SEQUENCE 26 AA; 2565 MW; 1D76859A7C45AF3F CRC64;

Query Match 19.0%; Score 38; DB 12; Length 26;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 TFPHSGSHHTH 19
| : | : |||
DB 2 TYTGGSAHTH 13

RESULT 12
Q91ZX5 PRELIMINARY; PRT; 27 AA.
AC Q91ZX5;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]
DE (FRAGMENT).
DE GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2;
RA Ross S., Viazov S., Roggendorf M.;
RT "An outbreak of Hepatitis C in a municipal hospital: A unique case of
RT patient-to-provider-to-patient transmission.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NS1 IS
CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY
CC SIMILARITY).
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1
CC FAMILY.
CC EMBL; AF227765; AAF72125.1; -.
DR InterPro: IPR002531; HCV_NSI.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27

```

```

SQ SEQUENCE 27 AA; 2808 MW; DFED6FAE38A8BCD4 CRC64;

Query Match 19.0%; Score 38; DB 12; Length 27;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 10 HPSGG--SHHTH 19
| | | | : |||
DB 3 HVSGGNVAHTH 14

RESULT 13
Q37931 PRELIMINARY; PRT; 28 AA.
ID Q37931
AC Q37931;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE GENE 3 (FRAGMENT).
DE GENE 3.
OS Bacteriophage Ike.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=10867;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97169154; PubMed=9016961;
RA Bruno R., Bradbury A.;
RT "A natural longer glycine-rich region in Ike filamentous phage confers
RT no selective advantage.";
RL Gene 184:121-123(1997).
DR EMBL; X95633; CAA64886.1; -.
FT NON_TER 1
FT NON_TER 28
FT NON_TER 28
SQ SEQUENCE 28 AA; 2074 MW; AE01C5995681DE45 CRC64;

Query Match 19.0%; Score 38; DB 9; Length 28;
Best Local Similarity 40.7%; Pred. No. 2.1e+02;
Matches 11; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 DTFHPGSGSHHTHSEDCGSHGSGEGG 33
| | | | | | | | | | | |
DB 1 DTGGGSGTGGDTGGSGGGSGGGSGG 27

RESULT 14
Q78381 PRELIMINARY; PRT; 17 AA.
ID Q78381
AC Q78381;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE VIRAL SAMPLE FLPBRSF (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION
DE (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.O., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
RA Korber B.T.M., Mullins J.I., Schochetman G., Berkelman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
RA Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92126; AAA44496.1; -.
FT NON_TER 1
FT NON_TER 1

```

```
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 1708 MW; 347570D2D12CA370 CRC64;

Query Match 18.5%; Score 37; DB 12; Length 17;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 TRDQDTFHPSGG 14
   | : || | |
Db 5 TNGSEIFRPGCG 16

RESULT 15
Q78345
ID Q78345 PRELIMINARY; PRT; 17 AA.
AC Q78345;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE VIRAL SAMPLE FLPAR5F (FLORIDA PATIENT A), PARTIAL ENV CDS, V5 REGION
DE (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.Q., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.;
RA Korber B.T.M., Mullins J.I., Schochetman G., Berkelman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
RA Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92114; AAA44470.1; -.
FT NON_TER 1 1
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 1635 MW; 3E83A0BFD12CA370 CRC64;

Query Match 18.0%; Score 36; DB 12; Length 17;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TRDQDTFHPSGG 14
   | : || | |
Db 5 TNGSEIFRPGCG 16
```

Search completed: February 21, 2002, 16:55:59  
Job time: 183 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:44:21 ; Search time 23.87 seconds  
(without alignments)  
108.612 Million cell updates/sec

Title: US-08-753-851-8  
Perfect score: 200  
Sequence: 1 CATRDQDTFHPSGSHTTHESEDPGSHSQEGGAN 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 253143

Minimum DB seq length: 0  
Maximum DB seq length: 35

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	63.0	23	15 AAR53480	CD44 peptide CD44-
2	96	48.0	20	20 AAY15202	Surrounding sequen
3	47	23.5	29	22 AAM14904	Peptide #1338 enco
4	47	23.5	29	22 AAM27333	Peptide #1370 enco
5	47	23.5	29	22 AAM02629	Peptide #1311 enco
6	45	22.5	35	22 AAM33720	Peptide #7757 enco
7	43	21.5	15	20 AAY15206	Surrounding sequen
8	43	21.5	34	20 AAW37649	Myc-His tag. Synt
9	42	21.0	18	16 AAR97264	Ubiquitin peptide
10	41	20.5	24	22 AAE06626	Human ANC_2H01 pro
11	41	20.5	24	22 AAB49884	Human zinc finger

12	40.5	20.2	20	17 AAW07626	Human high polymer
13	40.5	20.2	21	21 AAY93344	Light chain of hum
14	40	20.0	24	12 AAW14821	myc oncogene prote
15	40	20.0	24	21 AAY52624	c-myc encoded onco
16	40	20.0	26	20 AAY09048	BSP-GST 5' deletio
17	39	19.5	31	18 AAW36465	Cyclitol-ubiquinon
18	39	19.5	34	22 AAG76858	Human colon cancer
19	38.5	19.2	14	22 AAB62422	Human SAPL peptide
20	38.5	19.2	35	18 AAW17096	Phage-derived flex
21	38	19.0	27	22 AAM13664	Peptide #98 encode
22	38	19.0	27	22 AAM26063	Peptide #100 encod
23	38	19.0	27	22 AAM01414	Peptide #96 encode
24	38	19.0	30	18 AAW36466	Cyclitol-ubiquinon
25	38	19.0	32	21 AAY44209	Fusion protein com
26	38	19.0	33	21 AAB08789	Conserved sequence
27	37	18.5	10	22 AAG97418	Human complementar
28	37	18.5	25	20 AAY32946	Mutant TD regulato
29	37	18.5	25	20 AAY05710	Threonine dehydrat
30	37	18.5	28	21 AAY81597	Human high molecu
31	37	18.5	29	20 AAY25326	Human pancreatic P
32	37	18.5	34	21 AAB12350	Fragment of human
33	36.5	18.2	32	22 AAB98752	Human secondary si
34	36.5	18.2	33	22 AAM17209	Peptide #3643 enco
35	36.5	18.2	33	22 AAM29704	Peptide #3741 enco
36	36.5	18.2	33	22 AAM04906	Peptide #3588 enco
37	36	18.0	12	21 AAY81994	Human high molecu
38	36	18.0	15	15 AAR60513	Hexahistidine-cont
39	36	18.0	16	21 AAY51161	Modified vpl pepti
40	36	18.0	16	21 AAY76261	Fragment of human
41	36	18.0	20	17 AAW08049	HIV peptide #34.
42	36	18.0	22	15 AAR53479	CD44 peptide CD44-
43	36	18.0	25	20 AAY76576	Human ovarian tumo
44	36	18.0	28	22 AAM15382	Peptide #1816 enco
45	36	18.0	28	22 AAM27867	Peptide #1904 enco

#### ALIGNMENTS

RESULT 1  
AAR53480  
ID AAR53480 standard; peptide; 23 AA.  
XX  
AC AAR53480;  
XX  
DF 01-DEC-1994 (first entry)  
XX  
DE CD44 peptide CD44-7.  
XX  
KW Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autoimmune disease;  
KW immunosuppressive.  
XX  
OS Homo sapiens.  
XX  
PN WO9409811-A.  
XX  
PD 11-MAY-1994.  
XX  
PF 29-OCT-1993; 93WO-US10412.  
XX  
PR 30-OCT-1992; 92US-0973339.  
XX  
(UYDU-) UNIV DUKE.  
XX  
Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
XX  
WPI; 1994-167121/20.  
XX  
Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 221-255 of the CD44 protein sequence.

XX Sequence 23 AA;

Query Match 63.0%; Score 126; DB 15; Length 23;

Best Local Similarity 100.0%; Pred. No. 4.3e-11;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATRDQDTFHPGSGSHTHSE 22

Db 1 catrdqdtfhpqsgshthse 22

RESULT 2

AA15202

ID AAY15202 standard; Peptide; 20 AA.

XX AC AAY15202;

DT 26-OCT-1999 (first entry)

DE Surrounding sequence of potential assembly site E15-Rg, in CD44.

KW isoform: CD44; expression vector; exon; modification;

KW proteoglycan; glycosaminoglycan binding protein; rheumatoid;

KW arthritis; asthma; immunological disorder; assembly site.

OS Synthetic.

OS Homo sapiens.

XX WO9337317-A1.

PN 29-JUL-1999.

PD 21-JAN-1999; 99WO-US01411.

PF 24-JAN-1998; 98US-0072416.

PR (BRIM ) BRISTOL-MYERS SQUIBB CO.

PA Aruffo AA, Bennett KL, Greenfield WB, Wolff EA;

XX WPI; 1999-478982/40.

XX New artificial proteoglycans useful for treating rheumatoid

PT arthritis

XX Example 1; Page 30; 63pp; English.

XX This is part of the surrounding sequence of one of the potential

CC assembly sites in a CD44 exon.

CC The recombinant fusion protein may be administered to a patient to

CC enhance the biological activity of a GAG-binding protein within the body.

CC The protein also enhances wound healing and may be used to treat

CC rheumatoid arthritis, asthma, chronic obstructive pulmonary disorder,

CC Lupus, inflammatory bowel disease, psoriasis, osteoarthritis or

CC HIV infection.

CC The fusion protein may be used to target proteoglycans to a given site,

CC which causes local accumulation of GAG-binding proteins. Therefore the

CC protein may be used as adjuvants for vaccination and in the targeting of

CC chemokines to non-immunogenic tumour cells to enhance cellular antitumour

CC responses. In addition, the fusion protein can also enhance the half-life

XX of non-CAG binding growth factors.

XX Sequence 20 AA;

Query Match 48.0%; Score 96; DB 20; Length 20;

Best Local Similarity 94.4%; Pred. No. 5.9e-07;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 DQDTFHPGSGSHTHSE 22

Db 1 dqdtfhpqsgshthgse 18

RESULT 3

AAM14904

ID AAM14904 standard; Protein; 29 AA.

XX AC AAM14904;

DT 12-OCT-2001 (first entry)

DE Peptide #1338 encoded by probe for measuring cervical gene expression.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer.

XX OS Homo sapiens.

XX WO200157278-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00670.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

PS Claim 27; SEQ ID No 19730; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes

CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

CC can be used to produce a single exon microarray, which can be used for

CC measuring human gene expression in a sample derived from human cervical

CC epithelial cells. By measuring gene expression, the probes are therefore

CC useful in grading and/or staging of diseases of the cervix, notably

CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 29 AA;

Query Match 23.5%; Score 47; DB 22; Length 29;

Best Local Similarity 42.9%; Pred. No. 7.3;

Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 11 PSGSGSHTHSEDDGSHSQE 31

Db 2 pspqhkshlenqqphngsg 22

```

RESULT 4
ID AAM27333 standard; Protein; 29 AA.
XX AC AAM27333;
XX AC AAM27333;
DT 17-OCT-2001 (first entry)
XX Peptide #1370 encoded by probe for measuring placental gene expression.
DE Probe: microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
KW Homo sapiens.
XX WO200157272-A2.
XX PN Penn SG, Hanzel DK, Chen W, Rank DR;
PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00663.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX Claim 27; SEQ ID No 27602; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.
XX SQ Sequence 29 AA;
XX Query Match 23.5%; Score 47; DB 22; Length 29;
XX Best Local Similarity 42.9%; Pred. No. 7.3;
XX Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
XX
XX QY 11 PSGGSHTTHESEDGSHGSOE 31
XX || | : | : | : | : |
XX 2 pspqhkkslenqpphngsg 22
XX
XX Db
XX
XX RESULT 5
XX AAM02629
XX ID AAM02629 standard; Protein; 29 AA.
XX AC AAM02629;
XX AC AAM02629;
XX 09-OCT-2001 (first entry)
XX Peptide #1311 encoded by probe for measuring breast gene expression.
XX Probe; human; breast disease; breast cancer; development disorder;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

```

```

XX Homo sapiens.
XX WO200157270-A2.
XX 09-AUG-2001.
XX 29-JAN-2001; 2001WO-US00661.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX Claim 27; SEQ ID No 11369; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes
XX (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for measuring human gene expression in
XX a human breast sample, where the probe hybridises at high stringency to a
XX nucleic acid expressed in the human breast. The probes are useful for
XX predicting, diagnosing, grading, staging, monitoring and prognosing
XX diseases of the human breast, particularly those diseases with polygenic
XX aetiology. The diseases include: breast cancer, disorders of development,
XX inflammatory diseases of the breast, fibrocystic changes, proliferative
XX breast disease and non-carcinoma tumours.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 29 AA;
XX Query Match 23.5%; Score 47; DB 22; Length 29;
XX Best Local Similarity 42.9%; Pred. No. 7.3;
XX Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
XX
XX QY 11 PSGGSHTTHESEDGSHGSOE 31
XX || | : | : | : | : |
XX 2 pspqhkkslenqpphngsg 22
XX
XX Db
XX
XX RESULT 6
XX AAM33720
XX ID AAM33720 standard; Protein; 35 AA.
XX AC AAM33720;
XX AC AAM33720;
XX 17-OCT-2001 (first entry)
XX Peptide #7757 encoded by probe for measuring placental gene expression.
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX Homo sapiens.
XX WO200157272-A2.
XX 09-AUG-2001.
XX

```





CC fusion proteins. BS202 cDNA clones and polypeptides are useful  
 CC for developing products for detecting, diagnosing, staging,  
 CC preventing or treating diseases or conditions of the breast,  
 CC including breast cancer.

SQ Sequence 34 AA;

Query Match 21.5%; Score 43; DB 20; Length 34;  
 Best Local Similarity 43.8%; Pred. No. 32;  
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 12 GSGSHTTHESEDGHSH 27

Db 10 asasmthsthrghsh 25

RESULT 9

ID AAR67264 standard; peptide; 18 AA.

XX AAR67264;

DT 27-JUL-1995 (first entry)

DE Ubiquitin peptide extension, PEST4.

XX Ubiquitin; peptide extension; assay; protein kinase; C-terminal;  
 KW Ha-RAS protein; farnesyl-protein transferase;  
 KW carboxy methyl transferase.

XX Synthetic.

PN US5366871-A.

PD 22-NOV-1994.

PF 13-NOV-1991; 91US-0791935.

PR 13-NOV-1991; 91US-0791935.

XX (UTAH ) UNIV UTAH.

PI Rechsteiner MC, Yoo YJ;

DR WPI; 1995-005826/01.

XX Assay for enzymes that modify peptide chains - using a substrate  
 PT comprising a ubiquitin-peptide extension contg. a sequence  
 PT modified by the enzyme

PS Claim 23; Column 23; 14pp; English.

XX The sequences given in AAR67262-70 represent ubiquitin-peptide  
 CC extensions which are used in the method of the invention for  
 CC assaying enzymes that modify peptide chains. These sequences  
 CC represent extensions which are known to be modified in the presence  
 CC of an enzyme being assayed for, and the reaction mixture formed is  
 CC then analysed. This method may be used for assaying protein kinases  
 CC or enzymes which modify the C-terminal end of Ha-RAS protein, such  
 CC as farnesyl-protein transferase, or carboxy methyl transferase.  
 CC These peptides are less expensive to synthesise than standard  
 CC peptides used in these methods and they may be purified from  
 CC bacterial extracts by simple acid extraction.

SQ Sequence 18 AA;

Query Match 21.0%; Score 42; DB 16; Length 18;  
 Best Local Similarity 47.1%; Pred. No. 20;  
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 13 GSGSHTTHESEDGHSHS 29

Db 2 ggmeafheseeeehhs 18

RESULT 10

AAE06626

ID AAE06626 standard; Protein; 24 AA.

XX AAE06626;

DT 25-SEP-2001 (first entry)

DE Human ANC\_2H01 protein zinc finger motif 5.

XX Human; ANC\_2H01 protein; catenin-binding protein; signal transduction;  
 KW gene regulation; zinc finger protein; alphaN-catenin; drug screening;  
 KW therapy; cancer; neurological disorder; cytostatic; neuroprotective.

XX Homo sapiens.

PN WO200147954-A2.

PD 05-JUL-2001.

PF 18-MAY-2000; 2000WO-EP04535.

PR 23-DEC-1999; 99EP-0204512.

XX (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Van Roy F, Vanlandschoot A, Janssens B;

XX WPI; 2001-418220/44.

XX Novel recombinant nucleic acids useful for diagnosing, prognosing  
 PT and/or treating cancer and neurological disorders, corresponds to a  
 PT protein binding to alpha-catenin protein and with signal transduction  
 PT function

PS Example 1; Fig 3; 160pp; English.

XX The invention relates to human catenin-binding proteins and their  
 CC corresponding cDNA molecules which functions in signal transduction  
 CC and gene regulatory pathways. The invention also provides an isolated  
 CC and/or recombinant nucleic acid or its functional fragment, homologue  
 CC or derivative, corresponding to a alpha-catenin binding protein. The  
 CC invention also relates to a novel human zinc finger protein binding  
 CC with a member of the a-catulin/vinculin family, preferably with a human  
 CC isoform of alpha N-catenin (neural form). The invention also relates to  
 CC the field of drug discovery, diagnosis, prognosis and treatment of  
 CC cancer and neurological disorders. The present sequence is human  
 CC ANC\_2H01 protein zinc finger motif.

SQ Sequence 24 AA;

Query Match 20.5%; Score 41; DB 22; Length 24;

Best Local Similarity 58.3%; Pred. No. 40;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 16 HTTTHSESDGHSH 27

Db 8 hetndpedlsh 19

RESULT 11

AAB49884

ID AAB49884 standard; Peptide; 24 AA.

XX AAB49884;

DT 07-MAR-2001 (first entry)

DE Human zinc finger motif ZF5.  
 XX  
 KW Catenin-binding zinc finger protein; cancer; neurological disorder;  
 KW drug screening.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1054059-A1.  
 XX  
 PD 22-NOV-2000.  
 XX  
 PF 17-MAY-1999; 99EP-0201543.  
 XX  
 PR 17-MAY-1999; 99EP-0201543.  
 XX  
 PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 XX  
 PI Van Roy F, Vanlandschoot A, Janssens B;  
 XX  
 DR WPI; 2001-033776/05.  
 XX  
 PT Nucleic acid or its fragments, useful for diagnosing and treating  
 PT cancer and neurological disorders, corresponds to a catenin-binding  
 PT protein in signal transduction and gene regulatory pathways -  
 XX  
 PS Disclosure; Page 52; 71pp; English.  
 XX  
 CC The present invention is related to the coding sequence and protein  
 CC fragments of a human catenin-binding zinc finger protein. The coding  
 CC sequence was isolated from a human kidney cDNA library, but is expressed  
 CC in most human tissue. The sequences provided by the invention can be used  
 CC in the diagnosis and treatment of cancer and neurological disorders, and  
 CC in drug screening to identify compounds capable of the same.  
 XX  
 SQ Sequence 24 AA;  
 XX  
 Query Match 20.5%; Score 41; DB 22; Length 24;  
 Best Local Similarity 58.3%; Pred. No. 40;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 16 HTTHESEGDGHSH 27  
 Db 8 hetndpedlhsh 19  
 XX  
 RESULT 12  
 AAW07626  
 ID AAW07626 standard; peptide; 20 AA.  
 XX  
 AC AAW07626;  
 XX  
 DT 04-FEB-1997 (first entry)  
 XX  
 DE Human high polymer quininogen L-chain derived peptide.  
 XX  
 KW Human; high polymer; quininogen; L-chain; cell adhesion;  
 KW cancer metastasis; platelet aggregation; inhibition; wound;  
 KW inflammatory disease; arteriosclerosis; glomerular nephritis;  
 KW treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..19  
 FT /note= "claimed peptide (claim 1)"  
 XX  
 PN JP08208692-A.  
 XX  
 PD 13AUG-1996.  
 XX  
 PF 28-SEP-1995; 95JP-0276418.  
 XX

PR 28-SEP-1994; 94JP-0259451.  
 XX  
 PA (SUMU ) SUMITOMO SEIYAKU KK.  
 XX  
 DR WPI; 1996-421988/42.  
 XX  
 PT Cell adhesion inhibiting peptide(s), used as cancer metastasis  
 PT inhibitor - comprises partial amino acid sequence of human high  
 PT polymer quininogen L chain  
 XX  
 PS Claim 3; Page 2; 14pp; Japanese.  
 XX  
 CC The present peptide, and its claimed fragment, are derived from  
 CC residues 402-498 of the human high polymer quininogen L-chain. They  
 CC are useful in cell adhesion, cancer metastasis or platelet  
 CC aggregation inhibitors, and in wound, inflammatory disease,  
 CC arteriosclerosis or glomerular nephritis treating agents. The  
 CC present peptide was synthesised using a solid phase method, and  
 CC purified using a YMC-DOS-120A-S15/13 column.  
 XX  
 SQ Sequence 20 AA;  
 XX  
 Query Match 20.2%; Score 40.5; DB 17; Length 20;  
 Best Local Similarity 44.4%; Pred. No. 38;  
 Matches 8; Conservative 1; Mismatches 8; Indels 1; Gaps 1;  
 QY 10 HPSGGSHTHSESDGHSH 27  
 Db 1 hnlghgh-kherdghgh 17  
 XX  
 RESULT 13  
 AAY93344  
 ID AAY93344 standard; peptide; 21 AA.  
 XX  
 AC AAY93344;  
 XX  
 DT 04-SEP-2000 (first entry)  
 XX  
 DE Light chain of human high molecular weight kininogen fragment.  
 XX  
 KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;  
 KW plasma kallikrein; heavy chain; light chain; analogue; angiotensin;  
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO200027415-A2.  
 XX  
 PD 18-MAY-2000.  
 XX  
 PF 09-NOV-1999; 99WO-US26377.  
 XX  
 PR 10-NOV-1998; 98US-0107844.  
 XX  
 PA (UTEM ) UNIV TEMPLE.  
 PA (DUPO ) DUPONT PHARM CO.  
 PA (COLM/) COLMAN W R.  
 PA (MOUS/) MOUSA A S.  
 XX  
 PI Colman WR, Mousa AS;  
 XX  
 DR WPI; 2000-376306/32.  
 XX  
 PT Method for inhibiting endothelial cell proliferation, using compound  
 PT that inhibit endothelial cell migration -  
 XX  
 PS Claim 3; Page 35; 41pp; English.  
 XX  
 CC The present sequence represents a fragment of the light chain of human  
 CC high molecular weight kininogen. It is used to produce compounds of

the invention. High molecular weight kininogen is a 120 kDa glycoprotein which binds with high affinity to endothelial cells, where it is cleaved by plasma kallikrein into heavy and light chains. Analogues of high molecular weight kininogen are used in the method of the invention. The specification describes a method of inhibiting endothelial cell proliferation. The method comprises contacting endothelial cells with a compound containing high molecular weight kininogen analogues. The method and the compounds can be used for inhibiting endothelial cell proliferation. The compounds can also be used for inhibiting angiogenesis. The compounds can also be used to inhibit migration of endothelial cells to vitronectin.

Sequence 21 AA;

Query Match 20.2%; Score 40.5; DB 21; Length 21;  
Best Local Similarity 44.4%; Pred. NO. 40;  
Matches 8; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 10 HPSGGSHTHSEDEGHSH 27

Db 2 hnlghgh-kherdqghgh 18

RESULT 14

AAW14821  
ID AAW14821 standard; peptide; 24 AA.

AC AAW14821;

DT 23-MAY-1997 (first entry)

XX myc oncogene protein residues 208-231.

XX Oncogene; monoclonal receptor; antibody; immunoglobulin; ligand;  
KW immunogen; epitope; oncoprotein; detection.

XX Homo sapiens.

XX US5030565-A.

XX 09-JUL-1991.

XX 15-FEB-1985; 85US-0701954.

XX 16-APR-1987; 87US-0039534.

XX 17-AUG-1983; 83US-0524084.

XX 17-AUG-1984; 84US-0001304.

XX 15-FEB-1985; 85US-0701954.

XX 17-AUG-1984; 84WO-US01304.

XX 18-APR-1991; 91US-0687710.

XX 17-DEC-1993; 93US-0170649.

XX (SCRI ) SCRIPPS RES INST.

XX Lerner RA, Niman HL;

XX WPI; 1991-222277/30.

XX Monoclonal receptors to protein, esp. onco-protein ligands - prepd.  
XX using a polypeptide corresp. to a portion of the protein aminoacid  
XX sequence

XX Disclosure; Page -: 4lpp; English.

XX The sequences given in AAW14803-32 represent peptides derived from  
XX oncogenes which are bound by the monoclonal receptors of the invention.  
XX The monoclonal receptor molecules are immunoglobulins which bind to  
XX both (a) a protein ligand and (b) a polypeptide having an amino acid  
XX residue sequence containing 7-40 amino acid residues corresponding to a  
XX sequence of a portion of the protein, the receptor molecule having been  
XX raised to an immunogen containing the polypeptide. High yields of  
XX monoclonal receptors can be obtained which bind to or immunoreact with

CC known predetermined epitopes of protein molecules such as oncoproteins.  
CC The receptors can be used for e.g. detection of oncoprotein ligands or  
CC in affinity sorbants for binding and purifying oncoprotein ligands.  
CC This sequence is specifically derived from normal human cells.

XX Sequence 24 AA;

Query Match 20.0%; Score 40; DB 12; Length 24;

Best Local Similarity 46.7%; Pred. No. 56;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATRDQDTHPSGGS 15

Db 1 casqssafspssds 15

RESULT 15

AAW52624  
ID AAW52624 standard; peptide; 24 AA.

XX AC AAW52624;

XX 28-FEB-2000 (first entry)

XX c-myc encoded oncoprotein epitope #24.

XX Oncoprotein; epitope; oncogene; retroviral; infection; cellular;  
XX monoclonal; antibody; MAb; purification; cancer; tumour; growth factor;  
KW mitogenic; expression; detection; diagnosis; prognosis; immunoassay;  
KW growth; development; neoplasia; foetus; non-invasive.  
KW oncoprotein.

XX Synthetic.

XX Homo sapiens.

XX US5985587-A.

XX 16-NOV-1999.

XX 02-JUN-1995; 95US-0461584.

XX 02-SEP-1994; 94US-0300068.

XX 21-MAY-1985; 85US-0736545.

XX 07-OCT-1991; 91US-0772702.

XX 17-AUG-1984; 84WO-US01304.

XX 15-FEB-1985; 85US-0702954.

XX (SCRI ) SCRIPPS RES INST.

XX Lerner RA, Niman HL;

XX WPI; 2000-022278/02.

XX Purifying oncoprotein ligands using monoclonal antibodies, useful for  
XX diagnosing cancers caused by retroviruses -

XX Claim 5; Columns 25-26; 52pp; English.

XX Sequences AAW52601-Y52675 represent oncoprotein epitopes used to raise  
XX monoclonal antibodies which bind to both the epitopes and the proteins  
XX that comprise them. Certain retroviruses are able to cause the formation  
XX of solid tumours within a short period of time after infection of the  
XX host. Oncogenes, and the oncoproteins they encode, are responsible for  
XX the tumorigenic potential of these retroviruses. Retroviral oncogenes  
XX are closely related to and are derived from cellular oncogenes, which  
XX encode proteins with mitogenic activity such as growth factors. The  
XX invention relates to monoclonal anti-oncoprotein antibodies, and the  
XX method used to purify them. The method of the invention may be used for  
XX obtaining purified oncoprotein ligands from aqueous solutions. It may be  
XX used in this way to detect proteins produced in tumour cells to diagnose  
XX cancers caused by retroviruses. It may also be used for the  
XX prognostication of foetal development (and other growth states including

CC neoplasia) using either urine or other body fluid obtained by non  
CC invasive methods, the antibodies being used to assay for oncoprotein. As  
CC the antibodies bind to epitopes of known amino acid sequence, the type of  
CC oncoprotein being expressed in the patient may be determined.

XX  
SQ

Sequence 24 AA;

Query Match 20.0%; Score 40; DB 21; Length 24;

Best Local Similarity 46.7%; Pred. No. 56;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATRDQDTFHPSCGS 15

||::| | | |

Db 1 casqdsafspssds 15

Search completed: February 21, 2002, 16:52:51  
Job time: 510 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:50:06 ; Search time 12.58 seconds  
(without alignments)  
62.609 Million cell updates/sec

Title: US-08-753-851-8  
Perfect score: 200  
Sequence: 1 CATRDQDTFPHSGSHTTHESEDGSHSGEGGAN 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 139820

Minimum DB seq length: 0  
Maximum DB seq length: 35

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : 1: Issued Patents AA.\*  
2: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pap.\*  
7: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	21.0	18	1	US-07-791-935B-3
2	40.5	20.2	21	4	US-09-612-126-3
3	38	19.0	26	3	US-08-785-636-1
4	38	19.0	26	3	US-09-095-407-2
5	38	19.0	33	1	US-07-859-923A-20
6	38	19.0	33	2	US-08-484-107-20
7	36	18.0	15	1	US-08-469-486-48
8	36	18.0	15	2	US-08-469-658-48
9	36	18.0	20	1	US-08-218-025A-69
10	36	18.0	20	2	US-08-143-311B-7
11	35	17.5	18	1	US-08-001-907-1
12	35	17.5	18	1	US-08-250-309-1
13	35	17.5	18	1	US-08-343-305-1
14	35	17.5	18	1	US-08-343-795-1
15	35	17.5	18	1	US-08-408-092-1
16	35	17.5	18	1	US-08-374-001-1
17	35	17.5	18	1	US-08-343-313-1
18	35	17.5	18	1	US-08-408-418-1
19	35	17.5	18	1	US-08-642-373-1
20	34	17.0	15	1	US-08-555-579-4
21	34	17.0	15	4	US-09-406-781-4
22	34	17.0	19	4	US-08-882-046-11
23	34	17.0	20	1	US-08-466-033-85
24	34	17.0	20	1	US-08-466-033-89
25	34	17.0	20	2	US-08-444-733-85
26	34	17.0	20	2	US-08-444-733-89
27	34	17.0	20	2	US-08-464-134-85

28 34 17.0 20 2 US-08-464-134-89 Sequence 89, Appl  
29 34 17.0 20 2 US-08-461-361-85 Sequence 85, Appl  
30 34 17.0 20 2 US-08-461-361-89 Sequence 89, Appl  
31 34 17.0 20 2 US-08-485-910-85 Sequence 85, Appl  
32 34 17.0 20 2 US-08-485-910-89 Sequence 89, Appl  
33 34 17.0 20 5 PCT-US95-06266-69 Sequence 69, Appl  
34 34 17.0 20 5 PCT-US95-06266-73 Sequence 73, Appl  
35 34 17.0 25 1 US-08-466-033-97 Sequence 97, Appl  
36 34 17.0 25 2 US-08-444-733-97 Sequence 97, Appl  
37 34 17.0 25 2 US-08-464-134-97 Sequence 97, Appl  
38 34 17.0 25 2 US-08-461-361-97 Sequence 97, Appl  
39 34 17.0 25 2 US-08-485-910-97 Sequence 97, Appl  
40 34 17.0 25 4 US-08-772-282-1 Sequence 1, Appl  
41 34 17.0 25 5 PCT-US95-06266-81 Sequence 81, Appl  
42 34 17.0 26 3 US-08-814-052-34 Sequence 34, Appl  
43 34 17.0 26 3 US-08-812-829-26 Sequence 26, Appl  
44 33 16.5 12 4 US-09-331-362-10 Sequence 10, Appl  
45 33 16.5 13 4 US-09-331-362-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-07-791-935B-3  
; Sequence 3, Application US/07791935B  
; Patent No. 5366871  
; GENERAL INFORMATION:  
; APPLICANT: RECHSTEINER, MARTIN C.  
; APPLICANT: YOO, YUNG JOON  
; TITLE OF INVENTION: OBIQUITIN-PEPTIDE EXTENSIONS AS  
; TITLE OF INVENTION: ENZYME SUBSTRATES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Thorpe, No. 5366871th & Western  
; STREET: 9035 South 700 East, Suite 200  
; CITY: Sandy  
; STATE: Utah  
; COUNTRY: USA  
; ZIP: 84070  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage  
; COMPUTER: compaq LTE/286  
; OPERATING SYSTEM: DOS 4.01  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07791,935B  
; FILING DATE: 19911113  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: none  
; FILING DATE: na  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Western, M. Wayne  
; REGISTRATION NUMBER: 22,788  
; REFERENCE/DOCKET NUMBER: T310  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (801) 566-6633  
; TELEFAX: (801) 566-0750  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acid residues  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; US-07-791-935B-3

Query Match 21.0%; Score 42; DB 1; Length 18;  
Best Local Similarity 47.1%; Pred. No. 9.9;  
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
QY 13 GGSHTTTHSESDGSHSGS 29  
|| ||||| |

Db 2 GGMEFMESEEEESS 18

## RESULT 2

US-09-612-126-3

; Sequence 3, Application US/09612126

; Patent No. 6284726

; GENERAL INFORMATION:

; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF

; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH

; FILE REFERENCE: 6056-258 CT1

; CURRENT APPLICATION NUMBER: US/09/612,126

; PRIOR FILING DATE: 2000-07-07

; PRIOR FILING DATE: 1998-11-10

; PRIOR APPLICATION NUMBER: PCT/US99/26377

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 21

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Human high

; OTHER INFORMATION: molecular weight kininogen light chain amino acids

; OTHER INFORMATION: Lys(420) through Gly(440)

US-09-612-126-3

## Query Match

Best Local Similarity 20.2%; Score 40.5; DB 4; Length 21;

Mismatches 8; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 10 HPSGGSHTTTHSEDDGHSH 27

Db 2 HNLGHGH-KHERDQGHGH 18

## RESULT 3

US-08-785-636-1

; Sequence 1, Application US/08785636

; Patent No. 6027942

; GENERAL INFORMATION:

; APPLICANT: Yip, Tai-Tung

; TITLE OF INVENTION: Method and Apparatus for Desorption and Ionization of

; FILE REFERENCE: Hutchens

; CURRENT APPLICATION NUMBER: US/08/785,636

; CURRENT FILING DATE: 1997-01-17

; EARLIER FILING DATE: 1993-05-28

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 26

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: peptide

US-08-785-636-1

## Query Match

Best Local Similarity 19.0%; Score 38; DB 3; Length 26;

Mismatches 8; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10 HPSGGSHTTTHSEDDGHSHSQEGG 33

Db 3 HPHGHHPHGHHPHGHHPHGHHPHG 26

## RESULT 4

US-09-095-407-2

; Sequence 2, Application US/09095407

; Patent No. 6124137

; GENERAL INFORMATION:

; APPLICANT: Hutchens, T. William

; Yip, Tai-Tung

; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski L.L.P.

; STREET: 1301 McKinney, Suite 5100

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77010-3095

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/095,407

; FILING DATE: 10-Jun-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/068,896

; FILING DATE: 28-MAY-1993

; APPLICATION NUMBER: WO PCT/US94/06064

; APPLICATION NUMBER: US 08/556,951

; FILING DATE: 27-NOV-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Paul, Thomas D.

; REGISTRATION NUMBER: 32,714

; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (713) 651-5325

; TELEFAX: (713) 651-5246

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; TYPE: amino acid

; LENGTH: 26 amino acids

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-095-407-2

Query Match 19.0%; Score 38; DB 3; Length 26;

Best Local Similarity 33.3%; Pred. No. 53;

Mismatches 8; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10 HPSGGSHTTTHSEDDGHSHSQEGG 33

Db 3 HPHGHHPHGHHPHGHHPHGHHPHG 26

## RESULT 5

US-07-859-923A-20

; Sequence 20, Application US/07859923A

; Patent No. 5444044

; GENERAL INFORMATION:

; APPLICANT: Jiang, Shibo

; APPLICANT: Lin, Kang

; APPLICANT: Neurath, A. Robert

; TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES AS

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

ADDRESSEE: The Lindsley F. Kimball  
ADDRESSEE: Research Institute of the  
ADDRESSEE: New York Blood Center  
STREET: 310 E. 67th Street  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10021  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 5.25 inch, 360 kb  
COMPUTER: IBM XT compatible  
OPERATING SYSTEM: MS DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/859,923A  
FILING DATE: 19920326  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: No. 5444044e  
FILING DATE: No. 5444044e  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 Amino Acids  
TYPE: AMINO ACID  
TOPOLOGY: Linear  
US-07-859-923A-20

Query Match 19.0%; Score 38; DB 1; Length 33;  
Best Local Similarity 38.9%; Pred. No. 70;  
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 16 HTTHSEDSHGSHSQEGG 33  
| | | | |  
Db 15 HLIEESQOQTKNEQEGG 32

RESULT 6  
US-08-484-107-20  
; Sequence 20, Application US/08484107  
; Patent No. 5840843  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Shibo  
; APPLICANT: Lin, Kang  
; APPLICANT: Neurath, A. Robert  
; TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES AS  
; TITLE OF INVENTION: INHIBITORS OF HIV-1  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Lindsley F. Kimball  
; ADDRESSEE: Research Institute of the  
; ADDRESSEE: New York Blood Center  
; STREET: 310 E. 67th Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10021  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 5.25 inch, 360 kb  
; COMPUTER: IBM XT compatible  
; OPERATING SYSTEM: MS DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,107  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/859,923  
; FILING DATE: March 26, 1992  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 33 Amino Acids  
; TYPE: Amino Acid

TOPOLOGY: Linear  
US-08-484-107-20

Query Match 19.0%; Score 38; DB 2; Length 33;  
Best Local Similarity 38.9%; Pred. No. 70;  
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 16 HTTHSEDSHGSHSQEGG 33  
| | | | |  
Db 15 HLIEESQOQTKNEQEGG 32

RESULT 7  
US-08-469-486-48  
; Sequence 48, Application US/08469486  
; Patent No. 5739281  
; GENERAL INFORMATION:  
; APPLICANT: Thoenes, Hans Christian  
; APPLICANT: Holtet, Thor Las  
; APPLICANT: Etzerodt, Michael  
; TITLE OF INVENTION: Improved method for the refolding of  
; TITLE OF INVENTION: proteins  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version  
; SOFTWARE: #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,486  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/192,060  
; FILING DATE: February 4, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 06363/002001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 542 5070  
; TELEFAX: 617 542 8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-469-486-48

Query Match 18.0%; Score 36; DB 1; Length 15;  
Best Local Similarity 52.6%; Pred. No. 52;  
Matches 10; Conservative 0; Mismatches 3; Indels 6; Gaps 2;

QY 14 GSHTTHSEDSHGSHSQEGG 32  
| | | | |  
Db 2 GSH--HH---HHHGSIEG 14

RESULT 8  
US-08-469-658-48

; Sequence 48, Application US/08469658  
; Patent No. 5917018  
; GENERAL INFORMATION:  
; APPLICANT: Th egersen, Hans Christian  
; APPLICANT: Holtet, Thor Las  
; APPLICANT: Etzerodt, Michael  
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF  
; PROTEINS  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,658  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/192,060  
; FILING DATE: February 4, 1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 06363/002002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 542 5070  
; TELEFAX: 617 542 8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-469-658-48

Query Match 18.0%; Score 36; DB 2; Length 15;  
Best Local Similarity 52.6%; Pred. No. 52;  
Matches 10; Conservative 0; Mismatches 3; Indels 6; Gaps 2;

QY 14 GSHTHSESDGSHGSGEG 32  
Db 2 GSH--HH----HHGSGIEG 14  
||| | | | | |  
||| | | | | |

RESULT 9  
US-08-218-025A-69  
; Sequence 69, Application US/08218025A  
; Patent No. 5556744  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, David B.  
; APPLICANT: Ugen, Kenneth E.  
; APPLICANT: Williams, William V.  
; TITLE OF INVENTION: Methods and Compositions for Diagnosing  
; and Treating Certain HIV Infected Patients  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: P.O. Box 457, 321 No. 5556744ristown Road  
; CITY: Spring House  
; STATE: Pennsylvania

; COUNTRY: U.S.A.  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/218,025A  
; FILING DATE: 24-MAR-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/891,451  
; FILING DATE: 29-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: WST33A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 540-9206  
; TELEFAX: (215) 540-5818  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-218-025A-69

Query Match 18.0%; Score 36; DB 1; Length 20;  
Best Local Similarity 50.0%; Pred. No. 73;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TRDQDTFHPGG 14  
Db 4 TNDTEFRPGG 15  
| | | | |  
| | | | |

RESULT 10  
US-08-143-311B-7  
; Sequence 7, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991



CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-143-311B-7

Query Match 18.0%; Score 36; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATRDQDT 8  
|||||||  
DB 14 ATRDQDT 20

RESULT 11  
US-08-001-907-1  
; Sequence 1, Application US/08001907  
; Patent No. 5393527  
; GENERAL INFORMATION:  
; APPLICANT: Mallick, Adrien  
; APPLICANT: Feindt, Hans H.  
; APPLICANT: Hahn, Gerald D.  
; TITLE OF INVENTION: Stabilized Microspheres and Methods of  
; PREPARATION  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Becton, Dickinson and Company  
; STREET: One Becton Drive  
; CITY: Franklin Lakes  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07417

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/001,907  
FILING DATE: 19930104

CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-2209/P-2539  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-001-907-1

Query Match 17.5%; Score 35; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 88;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 14 GSHTTHESEDGH 25  
|:| |:| |

DB 2 GAHHAAHADAH 13

RESULT 12  
US-08-250-309-1  
; Sequence 1, Application US/08250309  
; Patent No. 5478741  
; GENERAL INFORMATION:  
; APPLICANT: Feindt, Hans H.  
; APPLICANT: Hahn, Gerald D.  
; APPLICANT: Maret, Melissa  
; APPLICANT: Uithoven, Keith  
; TITLE OF INVENTION: Improved Antibodies to P. falciparum  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Becton Dickinson and Company  
; STREET: One Becton Drive  
; CITY: Franklin Lakes  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07417

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/250,309  
FILING DATE:  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/945,287  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-2251  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-250-309-1

Query Match 17.5%; Score 35; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 88;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 14 GSHTTHESEDGH 25  
|:| |:| |

DB 2 GAHHAAHADAH 13

RESULT 13  
US-08-343-305-1  
; Sequence 1, Application US/08343305  
; Patent No. 5580735  
; GENERAL INFORMATION:  
; APPLICANT: Mallick, Adrien  
; APPLICANT: Feindt, Hans H.  
; APPLICANT: Hahn, Gerald D.  
; TITLE OF INVENTION: Stabilized Microspheres and Methods of  
; PREPARATION  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Becton, Dickinson and Company  
; STREET: One Becton Drive  
; CITY: Franklin Lakes  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07417

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/343,305  
FILING DATE: 22-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,907  
FILING DATE: 04-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-2209/P-2539  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-343-305-1

Query Match 17.5%; Score 35; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 88;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 14 GSHTTHESEDGH 25  
|:| | | | |  
Db 2 GAHHAHAADAH 13

RESULT 14  
US-08-343-795-1  
Sequence 1, Application US/08343795  
Patent No. 5593843  
GENERAL INFORMATION:  
APPLICANT: Malick, Adrien  
APPLICANT: Feindt, Hans H.  
APPLICANT: Hahn, Gerald D.  
TITLE OF INVENTION: Stabilized Microspheres and Methods of Preparation  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Becton, Dickinson and Company  
STREET: One Becton Drive  
CITY: Franklin Lakes  
STATE: NJ  
COUNTRY: US  
ZIP: 07417  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/343,795  
FILING DATE: 22-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,907  
FILING DATE: 04-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-2209/P-2539  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-343-795-1  
Query Match 17.5%; Score 35; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 88;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 14 GSHTTHESEDGH 25  
|:| | | | |  
Db 2 GAHHAHAADAH 13  
RESULT 15  
US-08-408-092-1  
Sequence 1, Application US/08408092  
Patent No. 5604117  
GENERAL INFORMATION:  
APPLICANT: Feindt, Hans H.  
APPLICANT: Hahn, Gerald D.  
APPLICANT: Maret, Melissa  
APPLICANT: Uithoven, Keith  
TITLE OF INVENTION: Improved Antibodies to P. falciparum  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Becton Dickinson and Company  
STREET: One Becton Drive  
CITY: Franklin Lakes  
STATE: NJ  
COUNTRY: US  
ZIP: 07417  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/408,092  
FILING DATE: 21-MAR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/945,287  
FILING DATE: 11-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-2251  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-408-092-1

Query Match 17.5%; Score 35; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 88;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 14 GSHTTHESEDGH 25  
|:| | | | |  
Db 2 GAHHAHAADAH 13

Search completed: February 21, 2002, 16:53:10  
Job time: 184 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:56:02 ; Search time 12.77 Seconds  
(without alignments)  
77.547 Million cell updates/sec

Title: US-08-753-851-9  
Perfect score: 74  
Sequence: 1 CRDGIRYVOKGEY 13  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues.

Total number of hits satisfying chosen parameters: 1822

Minimum DB seq length: 0  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	31.1	9	2 PS0253	glycine cleavage s
2	23	31.1	10	2 A43405	6-phosphofructo-2-
3	21	28.4	12	2 PH0746	protein QA300025 -
4	21	28.4	12	2 PH0746	T-cell receptor be
5	20	27.0	9	2 B39841	dextranucrase (EC
6	20	27.0	12	2 S18722	matK protein - bee
7	19	25.7	12	2 A34858	proteinase E - bla
8	19	25.7	12	2 PH0771	T-cell receptor be
9	18	24.3	8	2 PH0803	T-cell receptor al
10	18	24.3	10	2 PH0245	Ig heavy chain CRD
11	18	24.3	11	2 A59146	conotoxin au5a - c
12	18	24.3	11	2 B59146	conotoxin au5b - c
13	18	24.3	12	2 S09082	proteasome chain 1
14	18	24.3	12	2 S23188	Z protein - guinea
15	17.5	23.6	13	2 S47373	T-cell antigen rec
16	17	23.0	6	2 A35890	RNA-directed DNA p
17	17	23.0	9	2 A44787	calliFERamide 10
18	17	23.0	9	2 B46250	alpha-adaptin - bo
19	17	23.0	10	2 S70251	nitrogenase (EC 1.
20	17	23.0	11	2 S21127	precorrin methyltr
21	17	23.0	11	2 PH0929	T-cell receptor be
22	17	23.0	11	2 PH0914	T-cell receptor be
23	17	23.0	12	2 PT0228	Ig heavy chain CDR
24	17	23.0	13	2 PT0331	Ig heavy chain CRD
25	17	23.0	13	2 S47365	T-cell antigen rec
26	17	23.0	13	2 S47371	T-cell antigen rec
27	17	23.0	13	2 S47372	T-cell antigen rec
28	17	23.0	13	2 S47380	T-cell antigen rec
29	17	23.0	13	2 S47390	T-cell antigen rec

30 17 23.0 13 2 S47392 T-cell antigen rec  
31 17 23.0 13 2 S54344 glyceraldehyde-3-p  
32 16 21.6 6 2 PC4392 whey glycoprotein  
33 16 21.6 8 2 S37141 rpsA protein - Erw  
34 16 21.6 8 2 PT0653 T-cell receptor be  
35 16 21.6 9 2 A28495 conopressin G - co  
36 16 21.6 9 2 S39040 lysine-conopressin  
37 16 21.6 10 2 PH0807 T-cell receptor al  
38 16 21.6 11 2 S66196 alcohol dehydrogen  
39 16 21.6 11 2 S19301 endo-1,4-beta-xyla  
40 16 21.6 11 2 S61797 T-cell-specific tr  
41 16 21.6 12 2 C49215 urease (EC 3.5.1.5  
42 16 21.6 12 2 S28215 glucan endo-1,3-be  
43 16 21.6 12 2 S25039 Ig heavy chain V r  
44 16 21.6 13 2 G22565 R-phycocerythrin ga  
45 16 21.6 13 2 S14316 photosystem I 9K c

## ALIGNMENTS

RESULT 1  
PS0253  
glycine cleavage system protein H - rice (strain Nihonbare) (fragment)  
N:Alternate names: glycine decarboxylase complex H protein  
C:Species: Oryza sativa (rice)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 03-Feb-1994  
C:Accession: PS0253  
R:Tsuigita, A.  
submitted to JIPID, April 1993  
A:Reference number: PS0206  
A:Accession: PS0253  
A:Molecule type: protein  
A:Residues: 1-9 <TSU>

Query Match 31.1%; Score 23; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 3 DGIRY 7  
DB 5 DGLKY 9

RESULT 2  
A43405  
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Mar-2000  
C:Accession: A43405  
R:Ventura, F.; Rosa, J.L.; Ambrosio, S.; Pilakis, S.J.; Bartrons, R.  
J. Biol. Chem. 267, 17939-17943, 1992  
A:Title: Bovine brain 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Evidence  
A:Reference number: A43405; MUID:92388154  
A:Accession: A43405  
A:Molecule type: protein  
A:Residues: 1-10 <VEN>  
C:Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; ph  
C:Keywords: phosphoric monoester hydrolase; phosphotransferase

Query Match 31.1%; Score 23; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 KGEY 13  
DB 7 KGEY 10

RESULT 3  
PA0030

protein OA300025 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 06-Jun-1997  
C:Accession: PA0030  
R:Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
submitted to JIPID, July 1994  
A:Description: Separation and characterization of Arabidopsis proteins by two-dimensional  
A:Reference number: PA0001  
A:Accession: PA0030  
A:Molecule type: protein  
A:Residues: 1-12 <RAM>  
A:Experimental source: seed  
C:Keywords: seed

Query Match 28.4%; Score 21; DB 2; Length 12;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 IRYVQGE 12  
| :|||  
Db 2 IEPIQGE 9

## RESULT 4

PH0746  
T-cell receptor beta chain (B28) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999  
C:Accession: PH0746  
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A:Title: T cell receptor genes in a series of class I major histocompatibility complex-I  
allelic exclusion and antigen-specific repertoire.  
A:Reference number: PH0746; MUID:92078846  
A:Accession: PH0746  
A:Molecule type: mRNA  
A:Residues: 1-12 <CAS>  
A:Cross-references: EMBL:X60837; NID:g50098; PIDN:CAA43230.1; PID:g50099  
A:Experimental source: T lymphocyte  
C:Keywords: T-cell receptor

Query Match 28.4%; Score 21; DB 2; Length 12;  
Best Local Similarity 44.4%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRDGIRYVQ 9  
| :|||  
Db 1 CASSSRVEQ 9

## RESULT 5

B39841  
dextranucrase (EC 2.4.1.5) - Streptococcus sobrinus (fragment)  
C:Species: Streptococcus sobrinus  
C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 23-Jun-1993  
C:Accession: B39841  
R:Mooser, G.; Hefta, S.A.; Paxton, R.J.; Shively, J.E.; Lee, T.D.  
J. Biol. Chem. 266, 8916-8922, 1991  
A:Title: Isolation and sequence of an active-site peptide containing a catalytic aspartate  
A:Reference number: A39841; MUID:91224988  
A:Accession: B39841  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <MOO>  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 27.0%; Score 20; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DGIR 6  
| :|||  
Db 1 DGVR 4

## RESULT 6

Sl8722  
matK protein - beechdrops plastid (fragment)  
C:Species: plastid Epifagus virginiana (beechdrops)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Feb-1995  
C:Accession: Sl8722  
R:Morden, C.W.; Wolfe, K.H.; dePamphilis, C.W.; Palmer, J.D.  
EMBO J. 10, 3281-3288, 1991  
A:Title: Plastid translation and transcription genes in a non-photosynthetic plant:  
A:Reference number: Sl7794; MUID:92007779

A:Accession: Sl8722  
A:Molecule type: DNA  
A:Residues: 1-12 <MOR>  
A:Cross-references: EMBL:X61798  
C:Genetics:  
A:Gene: matK  
A:Genome: plastid  
C:Keywords: plastid

Query Match 27.0%; Score 20; DB 2; Length 12;  
Best Local Similarity 60.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 IRYVQ 9  
| :|||  
Db 8 LRYLQ 12

## RESULT 7

A34858  
proteinase E - blacktail rattlesnake (fragment)  
C:Species: Crotalus molossus nigrescens (blacktail rattlesnake)  
C:Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #text\_change 18-Jun-1993  
C:Accession: A34858  
R:Ramirez, G.A.; Fletcher Jr., P.L.; Possani, L.D.  
Toxicon 28, 285-297, 1990  
A:Title: Characterization of the venom from Crotalus molossus nigrescens Gloyd (black  
A:Reference number: A34858; MUID:90260877  
A:Accession: A34858  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-12 <RAM>

Query Match 25.7%; Score 19; DB 2; Length 12;  
Best Local Similarity 75.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 RYVQ 9  
| :|||  
Db 4 RYVZ 7

## RESULT 8

PH0771  
T-cell receptor beta chain (PE5.1.1) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999  
C:Accession: PH0771  
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A:Title: T cell receptor genes in a series of class I major histocompatibility complex  
allelic exclusion and antigen-specific repertoire.  
A:Reference number: PH0746; MUID:92078846  
A:Accession: PH0771  
A:Molecule type: mRNA  
A:Residues: 1-12 <CAS>

A:Cross-references: EMBL:X60865; NID:g53624; PIDN:CAA43255.1; PID:g53625  
 A:Experimental source: T lymphocyte  
 C:Keywords: T-cell receptor

Query Match 25.7%; Score 19; DB 2; Length 12;  
 Best Local Similarity 33.3%; Pred. No. 3.6e+03;  
 Matches 3; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRDGIRYVQ 9  
 I I I  
 Db 1 CASSFQYEQ 9

RESULT 9  
 PH0803  
 T-cell receptor alpha chain (J2) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PH0803  
 R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
 J. Exp. Med. 174, 1371-1383, 1991  
 A:Title: T cell receptor genes in a series of class I major histocompatibility complex-I allelic exclusion and antigen-specific repertoire.  
 A:Reference number: PH0746; MUID:92078846  
 A:Accession: PH0803  
 A:Molecule type: mRNA  
 A:Residues: 1-8 <CAS>  
 A:Cross-references: EMBL:X60912  
 A:Experimental source: T lymphocyte  
 C:Keywords: T-cell receptor

Query Match 24.3%; Score 18; DB 2; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRDGIRY 7  
 I I I  
 Db 1 CAAGITF 7

RESULT 10  
 PT0245  
 Ig heavy chain CRD3 region (clone 2-103C) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0245  
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: PT0222; MUID:91108337  
 A:Accession: PT0245  
 A:Molecule type: DNA  
 A:Residues: 1-10 <VAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 24.3%; Score 18; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 4.5e+03;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RDGIRY 7  
 I I I  
 Db 5 RDGYNW 10

RESULT 11  
 A59146  
 conotoxin au5a - cone shell (Conus aulicus)  
 C:Species: Conus aulicus (court cone)  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: A59146  
 R:Walker, C.S.; Steel, D.; Jacobsen, R.B.; Lirazan, M.B.; Cruz, L.J.; Hooper, D.; She  
 J. Biol. Chem. 274, 30664-30671, 1999  
 A:Title: The T-superfamily of conotoxins.  
 A:Reference number: A59147; MUID:99452958  
 A:Accession: A59146  
 A:Molecule type: protein  
 A:Residues: 1-11 <WAL>  
 C:Keywords: amidated carboxyl end; toxin; venom  
 F:2-9,3-10/Disulfide bonds: #status experimental  
 F:11/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 24.3%; Score 18; DB 2; Length 11;  
 Best Local Similarity 57.1%; Pred. No. 5e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRDGIRY 7  
 I I I  
 Db 2 CCPFIRY 8

RESULT 12  
 B59146  
 conotoxin au5b - cone shell (Conus aulicus)  
 C:Species: Conus aulicus (court cone)  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: B59146  
 R:Walker, C.S.; Steel, D.; Jacobsen, R.B.; Lirazan, M.B.; Cruz, L.J.; Hooper, D.; She  
 J. Biol. Chem. 274, 30664-30671, 1999  
 A:Title: The T-superfamily of conotoxins.  
 A:Reference number: A59147; MUID:99452958  
 A:Accession: B59146  
 A:Molecule type: protein  
 A:Residues: 1-11 <WAL>  
 C:Keywords: amidated carboxyl end; toxin; venom  
 F:2-9,3-10/Disulfide bonds: #status experimental  
 F:11/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 24.3%; Score 18; DB 2; Length 11;  
 Best Local Similarity 57.1%; Pred. No. 5e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRDGIRY 7  
 I I I  
 Db 2 CCPVIRY 8

RESULT 13  
 S09082  
 proteasome chain 1 - rat (fragment)  
 N:Alternate names: multicatalytic proteinase chain 1  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Aug-1998  
 C:Accession: S09082  
 R:Lilley, K.S.; Davison, M.D.; Rivett, A.J.  
 FEBS Lett. 262, 327-329, 1990  
 A:Title: N-terminal sequence similarities between components of the multicatalytic pr  
 A:Reference number: S09082; MUID:90242957  
 A:Accession: S09082  
 A:Molecule type: protein  
 A:Residues: 1-12 <Lilf>  
 C:Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 24.3%; Score 18; DB 2; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 5.5e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RDGI 5  
 I I I  
 Db 9 KDGI 12

## RESULT 14

S23168  
Z protein - guinea pig (fragment)  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 19-Mar-1997 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: S23168  
R:Nicollier, M.; Roblin, S.; Cypriani, B.; Remy-Martin, J.P.; Adessi, G.L.  
Eur. J. Biochem. 205, 1137-1144, 1992  
A:Title: Purification and characterization of a binding protein related to the Z class of  
A:Reference number: S23168; MUID:92249319  
A:Accession: S23168  
A:Molecule type: protein  
A:Residues: 1-12 <NIC>  
A:Experimental source: liver  
C:Function:  
A:Description: binds dehydroepiandrosterone sulfate, estrone sulfate, oleic acid, cholesterol  
C:Keywords: liver; steroid binding

Query Match 24.3%; Score 18; DB 2; Length 12;

Best Local Similarity 100.0%; Pred. No. 5.5e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 GEY 13

|||

Db 5 GEY 7

## RESULT 15

S47373  
T-cell antigen receptor VJ junction beta chain - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S47373  
R:Lehner, P.J.  
submitted to the EMBL Data Library, August 1994  
A:Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T cell  
A:Reference number: S47355  
A:Accession: S47373  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-13 <LEH>  
A:Cross-references: EMBL:Z35672; NID:g527489; PIDN:CAA84741.1; PID:g527490  
C:Keywords: T-cell receptor

Query Match 23.6%; Score 17.5; DB 2; Length 13;

Best Local Similarity 38.5%; Pred. No. 7.2e+03;

Matches 5; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 CRDGTRYVQKGEY 13

|

||

Db 1 CASSIRSADE-EY 12

Search completed: February 21, 2002, 16:59:02

Job time: 180 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:57:07 ; Search time 10.22 Seconds

(without alignments)  
46.638 Million cell updates/sec

Title: US-08-753-851-9

Perfect score: 74

Sequence: 1 CRDGIRYVQGEY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 530

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	28.4	13	1 CXET_CONTE	P81755 conus texti
2	18	24.3	10	1 UXA6_CHLTR	P38007 chlamydia t
3	17	23.0	9	1 FARA_CALVO	P41865 calliphora
4	17	23.0	10	1 RLAA_MOUSE	P99027 mus musculus
5	16	21.6	5	1 FARP_ARTHR	P41853 artiposthi
6	16	21.6	8	1 RSL_ERWCH	P37985 erwinia chr
7	16	21.6	9	1 CONO_CONGE	P05486 conus geogr
8	15	20.3	7	1 FARP_ASCSU	P31890 ascaris suu
9	15	20.3	8	1 ALI7_CARMA	P81820 carcinus ma
10	15	20.3	8	1 FARP_ASCSU	P43171 ascaris suu
11	15	20.3	9	1 ALI1_CARMA	P81814 carcinus ma
12	15	20.3	9	1 OXYT_BURPE	P42995 bufo regula
13	15	20.3	9	1 OXYT_CYPCA	P23879 cyprinus ca
14	15	20.3	9	1 OXYT_RABIT	P32878 oryctolagus
15	15	20.3	9	1 OXYV_SQUAC	P43000 squalus aca
16	15	20.3	9	1 SAP_STOVA	P24047 stomopneute
17	15	20.3	10	1 PSBF_CAPAN	P03367 capsicum an
18	15	20.3	12	1 SOIS_BACSU	P80863 bacillus su
19	15	20.3	13	1 LIGE_TRAVE	P29012 trameces ve
20	14	18.9	7	1 FARP_SALTY	P42564 hirudo medi
21	14	18.9	9	1 RSL1_SALTY	O54296 salmonella
22	14	18.9	9	1 TKCL_CALVO	P41517 calliphora
23	14	18.9	9	1 TKL1_LOCMI	P16223 locusta mig
24	14	18.9	9	1 ULAD_HUMAN	P31929 homo sapien
25	14	18.9	10	1 AMPN_HELAM	P81731 helicoverpa
26	14	18.9	10	1 SP34_DICMU	P81535 dictyosteli
27	14	18.9	10	1 SP1_HALRO	Q10997 halocynthia
28	14	18.9	10	1 TKL2_LOCMI	P16224 locusta mig
29	14	18.9	10	1 TKL3_LOCMI	P30249 locusta mig
30	14	18.9	10	1 TKL4_LOCMI	P30250 locusta mig
31	14	18.9	10	1 TRP5_LEUMA	P81737 leucophaea
32	14	18.9	10	1 UHA3_HUMAN	P40930 homo sapien
33	14	18.9	11	1 PKC1_CARMO	P82684 carausius m

34	14	18.9	11	1 TKC2_CALVO	P41518 calliphora
35	14	18.9	12	1 CD11_LITXA	P56245 litoria xan
36	14	18.9	12	1 CD14_LITXA	P56246 litoria xan
37	14	18.9	12	1 CXAL_CONIM	P50983 conus imper
38	14	18.9	12	1 YZPY_ECOLI	P17776 escherichia
39	14	18.9	13	1 FIBA_CAVPO	P14445 cavia porce
40	14	18.9	13	1 UNO2_PINPS	P81667 pinus pinas
41	13	17.6	5	1 PRCT_PERAM	P01373 periplaneta
42	13	17.6	7	1 FARI_ASCSU	P31889 ascaris suu
43	13	17.6	7	1 FARI_PROCL	P38499 procambarus
44	13	17.6	7	1 FAR2_PROCL	P38498 procambarus
45	13	17.6	8	1 FAR3_HOMAM	P41486 homarus ame

## ALIGNMENTS

RESULT 1  
CXET\_CONTE STANDARD; PRT; 13 AA.  
AC P81755;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE EPSILON-COMOTOXIN TXIX.  
OS Conus textile (Cloth-of-gold cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=6494;  
RN [1]  
RP SEQUENCE, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY NMR.  
RC TISSUE=Venom;  
RX MEDLINE=99254114; PubMed=10318957;  
RA Rigby A.C., Lucas-Meunier E., Kallume D.E., Czerwicz E., Hambe B.,  
RA Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,  
RA Furie B.C., Furie B., Stenflo J.P.;  
RT "A conotoxin from Conus textile with unusual posttranslational  
RT modifications reduces presynaptic Ca2+ influx.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).  
CC -!- FUNCTION: CONOTOXIN WHICH ACTS AT PRESYNAPTIC MEMBRANES, BLOCKING  
CC THE CALCIUM CHANNELS.  
CC -!- PTM: O-GLYCAN CONSISTS OF THE DISACCHARIDE GAL-GALNAC.  
DR PDB; 1WCT; 08-JUN-99.  
KW Presynaptic neurotoxin; Calcium channel inhibitor; Venom; Vitamin K;  
KW Gamma-carboxyglutamic acid; Glycoprotein; Hydroxylation; Bromination;  
KW 3D-structure. 2 8  
FT DISULFID 3 9  
FT DISULFID 1 1  
FT MOD\_RES 4 4  
FT MOD\_RES 7 7  
FT MOD\_RES 13 13  
FT CARBOHYD 10 10  
FT CARBOHYD 13 AA; 1388 MW; 386C9E1C74FA378 CRC64;  
SQ SEQUENCE

Query Match 28.4%; Score 21; DB 1; Length 13;  
Best Local Similarity 75.0%; Pred. No. 9.4e+02;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRDG 4  
DB 3 CEDG 6

RESULT 2  
UXA6\_CHLTR STANDARD; PRT; 10 AA.  
AC P38007;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE FROM ELEMENTARY BODY (FRAGMENT).



```

OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE.
RC STRAIN=L2/434/BU;
RA Binl L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA Comanducci M., Christensen G., Birkelund S., Vretou E., Ratti G.,
RA Pallini V.;
RL Submitted (SEP-1994) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS; ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.44, ITS MW IS: 38.6 KDA.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1243 MW; DAD39A33304B5339 CRC64;

Query Match 24.3%; Score 18; DB 1; Length 10;
Best Local Similarity 40.0%; Pred. No. 2.3e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 RYVQK 10
DB 4 KYIKK 8

RESULT 3
FARA_CALVO STANDARD; PRT; 9 AA.
AC P41865;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CALLIFMRFAMIDE 10.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92119611; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated callifmrfamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR: A4787; A4787.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9
FT UNSURE 1 1
FT OR S OR A.
SQ SEQUENCE 9 AA; 1103 MW; 29D00699CAB40457 CRC64;

Query Match 23.0%; Score 17; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RGIRY 7
DB 4 RDMRF 9

RESULT 4
RLA2_MOUSE STANDARD; PRT; 10 AA.
AC P9027;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

```

DE 60S ACIDIC RIBOSOMAL PROTEIN P2 (FRAGMENT).
GN RPLP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA Cowthorne M.;
RL Submitted (AUG-1998) to the SWISS-PROT data bank.
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -!- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
DR SWISS-2DPAGE; P9027; MOUSE.
KW Ribosomal protein; Phosphorylation.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1186 MW; 07121E3B45BDC2DB CRC64;

Query Match 23.0%; Score 17; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IRYV 8
DB 1 MRYV 4

RESULT 5
FARP_ARTTR STANDARD; PRT; 5 AA.
AC P41853;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFAMIDS-LIKE NEUROPEPTIDE RYIRF-AMIDE.
OC Artiopesthia triangulata.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
OC Rhabditophora; Seriata; Tricladida; Terricola; Geoplanidae;
OC Arthurdendyus.
OX NCBI_TaxID=132421;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=94211927; PubMed=7909164;
RA Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;
RT "RYIRFamide: a turbellarian FMRFamide-related peptide (FARP).";
RL Regul. Pept. 50:37-43(1994).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 5 5
FT OR S OR A.
SQ SEQUENCE 5 AA; 754 MW; 69D4004B44600000 CRC64;

Query Match 21.6%; Score 16; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 RYV 9
DB 1 RYIR 4

RESULT 6
RSL_ERWCH STANDARD; PRT; 8 AA.
ID RSL_ERWCH
AC P37985;

```

01-OCT-1994 (Rel. 30, Created)  
01-OCT-1994 (Rel. 30, Last sequence update)  
01-NOV-1995 (Rel. 32, Last annotation update)  
30S RIBOSOMAL PROTEIN S1 (FRAGMENT).  
RPSA.  
OS Erwinia chrysanthemi.  
OC Bacteria; Proteobacteria; gamma subphylum; Enterobacteriaceae;  
OC Pectobacterium.  
OX NCBI\_TaxID=556;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3937;  
RA Doullie A., Toussaint A., Faelen M.;  
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE  
CC SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X74750; CAA52769.1; -;  
DR PIR; S37141; S37141.  
KW Ribosomal protein; Repeat; RNA-binding.  
FT NON\_TER 1  
SQ SEQUENCE 8 AA; 837 MW; 9E18733DC5B339CD CRC64;  
  
Query Match 21.6%; Score 16; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 10 KGE 12  
DB 6 KGE 8  
  
RESULT 7  
CONO\_CONGE STANDARD; PRT; 9 AA.  
AC P05486;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE LYS-CONOPRESSIN G.  
OS Conus geographus (Geography cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=6491;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=88058932; PubMed=3680228;  
RA Cruz L.J., de Santos V., Zafaralla G.C., Ramilo C.A., Zeikus R.D.,  
RA Gray W.R., Olivera B.M.;  
RT "Invertebrate vasopressin/oxytocin homologs. Characterization of  
RT peptides from Conus geographus and Conus straitus venoms.";  
RL J. Biol. Chem. 262:15821-15824(1987).  
RN [2]  
RP REVIEW.  
RX MEDLINE=89024586; PubMed=3052286;  
RA Gray W.R., Olivera B.M., Cruz L.J.;  
RT "Peptide toxins from venomous Conus snails.";  
RL Annu. Rev. Biochem. 57:665-700(1988).  
CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
DR PIR; A28495; A28495.  
DR Interpro; IPR000981; Neurhypophys\_horm.  
DR Pfam; PF00220; hormone4; 1.

DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
KW Hormone; Amidation.  
FT DISULFID 1 6  
FT MOD\_RES 9 9 AMIDATION.  
SQ SEQUENCE 9 AA; 1037 MW; D4FC276EB4540059 CRC64;  
  
Query Match 21.6%; Score 16; DB 1; Length 9;  
Best Local Similarity 57.1%; Pred. No. 1e+05;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 5 IRYVQKG 11  
DB 3 IRNCPKG 9  
  
RESULT 8  
FAR2\_ASCSU STANDARD; PRT; 7 AA.  
ID FAR2\_ASCSU  
AC P31890;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE FMRFAMIDE-LIKE NEUROPEPTIDE AF2  
OS Ascaris suum (pig roundworm) (Ascaris lumbricoides), and  
OS Panagrellus redivivus.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;  
OC Ascarididae; Ascaris.  
OX NCBI\_TaxID=6253, 6233;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=A.suum;  
RX MEDLINE=93324431; PubMed=8332542;  
RA Cowden C., Stretton A.O.W.;  
RT "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";  
RL Peptides 14:423-430(1993).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=P.redivivus;  
RX MEDLINE=95060998; PubMed=7970891;  
RA Maule A.G., Shaw C., Bowman J.W.;  
RT "The FMRFamide-like neuropeptide AF2 (Ascaris suum) is present in the  
RT free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";  
RL Parasitology 109:351-356(1994).  
CC -!- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.  
CC -!- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF  
CC GLANDS PARTICULARLY IN THE ANTERIOR REGIONS.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7  
SQ SEQUENCE 7 AA; 992 MW; 69D4073B5B11E350 CRC64;  
  
Query Match 20.3%; Score 15; DB 1; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 10 KGEY 13  
DB 1 KHEY 4  
  
RESULT 9  
ALI7\_CARMA STANDARD; PRT; 8 AA.  
ID ALI7\_CARMA  
AC P81820;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CARCINOSTATIN 17.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 8 8 AMIDATION (POTENTIAL).  
 SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 20.3%; Score 15; DB 1; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 GEY 13  
 Db 2 GQY 4

RESULT 10  
 FAR7\_ASCSU  
 ID FAR7\_ASCSU STANDARD; PRT; 8 AA.  
 AC P43171;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE AF7.  
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;  
 OC Ascarididae; Ascaris.  
 OX NCBI\_TaxID=6253;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95380362; PubMed=7651904;  
 RA Cowden C., Stretton A.O.W.;  
 RT "Eight novel FMRFamide-like neuropeptides isolated from the nematode  
 RT Ascaris suum.";  
 RL Peptides 16:491-500(1995).  
 CC -!- SIMILARITY: BELONGS TO THE FARF (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 963 MW; 9CD40059D417687D CRC64;

Query Match 20.3%; Score 15; DB 1; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 1e+05;  
 Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 GRYVQ 9  
 Db 2 GPRFIR 7

RESULT 11  
 AL11\_CARMA  
 ID AL11\_CARMA STANDARD; PRT; 9 AA.  
 AC P81814;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 11.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 927 MW; 832D79CDCB46D861 CRC64;

Query Match 20.3%; Score 15; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 GEY 13  
 Db 3 GQY 5

RESULT 12  
 OXYT\_BUFRE  
 ID OXYT\_BUFRE STANDARD; PRT; 9 AA.  
 AC P42995;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE SERITOCIN.  
 OS Bufo regularis (African toad).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufonidae;  
 OC Bufo.  
 OX NCBI\_TaxID=8390;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE= pituitary neurointermediate lobe;  
 RX MEDLINE=96059313; PubMed=7591488;  
 RA Chauvet J., Michel G., Ouedraogo Y., Chou J., Chait B.T., Acher R.;  
 RT "A new neurohypophysial peptide, seritocin ([ser5, ile8]-oxytocin),  
 RT identified in a dryness-resistant African toad, Bufo regularis.";  
 RL Int. J. Pept. Protein Res. 45:482-487(1995).  
 CC -!- FUNCTION: DEVIOD OF OXYTOCIC ACTIVITY.  
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR InterPro: IPR000981; Neurohypophys\_horm.  
 DR Pfam: PF00220; hormone4; 1.  
 DR PROSITE: PS00264; NEUROHYPOPHYS\_HORM; FALSE\_NEG.  
 KW Hormone; Amidation.  
 FT MOD\_RES 1 6 AMIDATION.  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 983 MW; 17FF476EA5A6D04B CRC64;

Query Match 20.3%; Score 15; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 YVQ 9  
 Db 2 YIQ 4

RESULT 13  
 OXYT\_CYPCA  
 ID OXYT\_CYPCA STANDARD; PRT; 9 AA.  
 AC P23879;

DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE VASOTOCIN.  
 OS Cyprinus carpio (Common carp), and Petromyzon marinus (Sea lamprey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.  
 OX NCBI\_TaxID=7962, 7757;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=C. carpio; TISSUE=Pituitary;  
 RA Acher R., Chauvet J., Chauvet M.-T., Crepy D.;  
 RT "Characterization of neurohypophyseal hormones from a fresh water bony  
 fish, the carp (Cyprinus carpio). Comparison with hormones from sea  
 water bony fishes.";  
 RL Comp. Biochem. Physiol. 14:245-254(1965).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=P. marinus; TISSUE=Pituitary;  
 RX MEDLINE=88225976; PubMed=3371648;  
 RA Lane T.F., Sower S.A., Kawauchi H.;  
 RT "Arginine vasotocin from the pituitary gland of the lamprey  
 (Petromyzon marinus): Isolation and amino acid sequence.";  
 RL Gen. Comp. Endocrinol. 70:152-157(1988).  
 CC -!- FUNCTION: ANTIDIURETIC HORMONE.  
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR PIR; B61364; B61364.  
 DR PIR; S06375; S06375.  
 DR InterPro: IPR000981; Neurohypophys\_horm.  
 DR Pfam: PF00220; hormone4; 1.  
 DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Amidation.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9  
 FT SEQUENCE 9 AA; 1053 MW; 17EB176EB456D04B CRC64;  
 SQ  
 Query Match 20.3%; Score 15; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 7 YVQ 9  
 Db 2 YIQ 4  
 RESULT 14  
 OXYT\_RABIT STANDARD; PRT; 9 AA.  
 AC P32878; P01188;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE OXYTOCIN (OXYTOCIN).  
 OS Oryctolagus cuniculus (Rabbit), Hippopotamus amphibius (Hippopotamus),  
 OS Balaeoptera physalus (Finback whale) (Common rorqual), and  
 OS Tachyglossus aculeatus (Australian echidna), and  
 OS Hydrolagus collei (Spotted ratfish) (Pacific ratfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986, 9833, 9770, 49271, 7873;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=Rabbit;  
 RX MEDLINE=72215060; PubMed=5150741;  
 RA Chauvet J., Chauvet M.-T., Acher R.;  
 RT "Evolution of neurohypophyseal hormones: isolation of active  
 principles from rabbits and rats.";  
 RL Biochimie 53:1099-1104(1971).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=H. amphibius;

RX MEDLINE=71232719; PubMed=5406007;  
 RA Ferguson D.R., Pickering B.T.;  
 RT "Arginine and lysine vasopressins in the hippocampus  
 neurohypophysis.";  
 RL Gen. Comp. Endocrinol. 13:425-429(1969).  
 RN [3]  
 RP SEQUENCE.  
 RC SPECIES=B. physalus;  
 RA Acher R., Chauvet J., Chauvet M.-T.;  
 RT "Isolation of finback whale oxytocin and vasopressin.";  
 RL Nature 201:191-192(1964).  
 RN [4]  
 RP SEQUENCE.  
 RC SPECIES=A. aculeatus;  
 RX MEDLINE=73223515; PubMed=4515919;  
 RA Acher R., Chauvet J., Chauvet M.-T.;  
 RT "Neurohypophyseal hormones and evolution of tetrapods.";  
 RL Nature New Biol. 244:124-126(1973).  
 RN [5]  
 RP SEQUENCE.  
 RC SPECIES=H. colliiei;  
 RX MEDLINE=70088110; PubMed=5366118;  
 RA Pickering B.T., Heller H.;  
 RT "Oxytocin as a neurohypophyseal hormone in the holoccephalian  
 elasmobranch fish, Hydrolagus collei.";  
 RL J. Endocrinol. 45:597-606(1969).  
 CC -!- FUNCTION: OXYTOCIN CAUSES CONTRACTION OF THE SMOOTH MUSCLE OF THE  
 UTERUS AND OF THE MAMMARY GLAND.  
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR PIR; A91466; A91466.  
 DR PIR; A92774; A92774.  
 DR PIR; A93147; A93147.  
 DR PIR; A93408; A93408.  
 DR PIR; B90667; B90667.  
 DR PDB; 1XV1; 15-OCT-90.  
 DR PDB; 1XY2; 15-OCT-90.  
 DR InterPro: IPR000981; Neurohypophys\_horm.  
 DR Pfam: PF00220; hormone4; 1.  
 DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Hypothalamus; Amidation; 3D-structure.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9  
 FT SEQUENCE 9 AA; 1010 MW; 17F8376EB456D04B CRC64;  
 SQ  
 Query Match 20.3%; Score 15; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 7 YVQ 9  
 Db 2 YIQ 4  
 RESULT 15  
 OXYV\_SQUAC STANDARD; PRT; 9 AA.  
 AC P43000;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE VALTPOCIN.  
 OS Squalus acanthias (Spiny dogfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.  
 OX NCBI\_TaxID=7797;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=73031727; PubMed=5083097;  
 RA Acher R., Chauvet J., Chauvet M.-T.;  
 RT "Phylogeny of the neurohypophyseal hormones. Two new active peptides  
 isolated from a cartilaginous fish, Squalus acanthias.";  
 RL Eur. J. Biochem. 29:12-19(1972).

[2]  
RN SEQUENCE  
RX MEDLINE=72128038; PubMed=4622083;  
RA Acher R., Chauvet J., Chauvet M.-T., Fontaine M.;  
RT "Identification of 2 new neurohypophyseal hormones, valitocin (Val8-  
RT oxytocin) and aspartocin (Asn4-oxytocin) in a selachian fish, the  
RT spiny dog-fish (Squalus acanthias).";  
RL C. R. Acad. Sci., D, Sci. Nat. 274:313-316(1972).  
CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
DR InterPro; IPR000981; Neurohypophys\_horm.  
DR Pfam; PF00220; hormone4; 1.  
DR PROSITE: PS00264; NEUROHYPOPHYS\_HORM; 1.  
KW Hormone; Amidation.  
FT DISULFID 1 6  
FT MOD\_RES 9 9  
SQ SEQUENCE 9 AA; 996 MW; 17EDD76EB456D04B CRC64;

Query Match 20.3%; Score 15; DB 1; Length 9;  
Best Local Similarity 66.7%; Pred. NO. 1e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 YVQ 9  
i:i  
Db 2 YIQ 4

Search completed: February 21, 2002, 17:00:08  
Job time: 181 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2002, 16:56:52 ; Search time 22.09 seconds  
(without alignments)  
86.081 Million cell updates/sec

Title: US-08-753-851-9

Perfect score: 74

Sequence: 1 CRDGIRYVQKGEY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 1983

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

1: SP\_TREMBL17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	28.4	9	2 Q9R7H9	Q9r7h9 haemophilus
2	21	28.4	12	2 Q50959	Q50959 neisseria g
3	19	25.7	13	4 Q16141	Q16141 homo sapien
4	18	24.3	10	6 Q9TR48	Q9tr48 bos taurus
5	18	24.3	12	8 Q9TR203	Q9tr203 bos taurus
6	18	24.3	13	2 Q9RFZ4	Q9rfz4 mycoplasma
7	17	23.0	8	6 Q9BF92	Q9bf92 tursiops tr
8	17	23.0	9	6 Q9XT05	Q9xt05 macropus ru
9	17	23.0	12	4 Q9UMR0	Q9umr0 homo sapien
10	17	23.0	12	12 Q9IF00	Q9if00 human adeno
11	17	23.0	12	12 Q9IEZ9	Q9iez9 human adeno
12	17	23.0	13	6 Q9TU76	Q9tu76 ovis aries
13	17	23.0	13	8 Q9MOK6	Q9mqk6 rupicapra r
14	17	23.0	13	11 Q9CU06	Q9cu06 mus musculus
15	16	21.6	8	6 Q9BFC3	Q9bfc3 didelphis m
16	16	21.6	8	6 Q9BFC2	Q9bfc2 macropus h
17	16	21.6	8	6 Q9BFC1	Q9bfc1 choloepus h
18	16	21.6	8	6 Q9BFC0	Q9bfc0 choloepus d
19	16	21.6	8	6 Q9BFB9	Q9bfb9 euphractus

## ALIGNMENTS

RESULT 1

Q9R7H9 ID Q9R7H9 PRELIMINARY; PRT; 9 AA.  
AC Q9R7H9;  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-MAY-2000 (TREMREL. 13, Last annotation update)  
DE LIPOPROTEIN (FRAGMENT).  
GN NLPD.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98083063; PubMed=9422600;  
RA Martin K., Morlin G., Smith A., Nordyke A., Eisenstark A., Golomb M.;  
RT "The tryptophanase gene cluster of Haemophilus influenzae type b;  
RT evidence for horizontal gene transfer.";  
RL J. Bacteriol. 180:107-118(1998).  
DR EMBL; AF003252; AAB96582.1; -.  
KW Lipoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 9 AA: 1152 MW; 35A017673B4412D7 CRC64;

Query Match 28.4%; Score 21; DB 2; Length 9;  
Best Local Similarity 37.5%; Pred. No. 4.7e+05;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 DGIRYVQK 10

Db 1 DPNRYLPR 8

RESULT 2

Q50959 ID Q50959 PRELIMINARY; PRT; 12 AA.  
AC Q50959;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)

DE OPAL1 PROTEIN (FRAGMENT).  
 OS Neisseria gonorrhoeae.  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=485;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MS11A;  
 RX MEDLINE=89210824; PubMed=2854063;  
 RA Taha M.K., So M., Seifert H.S., Billyard E., Marchal C.;  
 RT "Pilin expression in Neisseria gonorrhoeae is under both positive and  
 negative transcriptional control.";  
 RL EMBO J. 7:4367-4378(1988).  
 DR EMBL; X13965; CAB37342.1;  
 FT NON\_TER 1  
 SQ SEQUENCE 12 AA; 1436 MW; 9684516C16C87735 CRC64;

Query Match 28.4%; Score 21; DB 2; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GIRY 7  
 Db 7 GVRV 10  
 I:|

RESULT 3  
 Q16141 ID Q16141 PRELIMINARY; PRT; 13 AA.

AC Q16141;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE BETA PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94220089; PubMed=7545922;  
 RA Sadakane Y., Maeda K., Kuroda Y., Horii K.;  
 RT "Identification of mutations in DNA polymerase beta mRNAs from  
 RT patients with Werner syndrome.";  
 RL Biochem. Biophys. Res. Commun. 200:219-225(1994).  
 DR EMBL; S69873; AAD14051.1;  
 FT NON\_TER 1  
 SQ SEQUENCE 13 AA; 1568 MW; D88C62798C9542CD CRC64;

Query Match 25.7%; Score 19; DB 4; Length 13;  
 Best Local Similarity 36.4%; Pred. No. 6.7e+03;  
 Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 RDGIYVQKGE 12  
 Db 3 RDSAVYINTPE 13  
 I:|

RESULT 4

Q3TR48 ID Q3TR48 PRELIMINARY; PRT; 10 AA.

AC Q3TR48;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE AMPHOTERIN (FRAGMENT).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OX NCBI\_TaxID=9913;  
 RN [1]

RP SEQUENCE.  
 RX MEDLINE=96029671; PubMed=7592757;  
 RA Hori O., Brett J., Slattery T., Cao R., Zhang J., Chen J.X.,  
 RA Nagashima M., Lundh E.R., Vijay S., Nitecki D.;  
 RT "The receptor for advanced glycation end products (RAGE) is a cellular  
 RT binding site for amphoterin. Mediation of neurite outgrowth and co-  
 RL J. Biol. Chem. 270:25752-25761(1995).  
 DR HSSP; P07155; 1HMF.  
 SQ SEQUENCE 10 AA; 1050 MW; 23B89A286761EB18 CRC64;

Query Match 24.3%; Score 18; DB 6; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 7.4e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KGEY 13  
 Db 2 KGEH 5  
 I:|

RESULT 5  
 Q9T2U3 ID Q9T2U3 PRELIMINARY; PRT; 12 AA.

AC Q9T2U3;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE NADH:UBIQUINONE OXIDOREDUCTASE (COMPLEX I) IRON-SULFUR PROTEIN  
 DE FRACTION 20 KDA POLYPEPTIDE PEPTIDE T-8.  
 OS Bos taurus (Bovine).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92138662; PubMed=1778979;  
 RA Masui R., Wakabayashi S., Matsubara H., Hatemi Y.;  
 RT "The amino acid sequence of the 9 kDa polypeptide and partial amino  
 RT acid sequence of the 20 kDa polypeptide of mitochondrial  
 RT NADH:ubiquinone oxidoreductase.";  
 RL J. Biochem. 110:575-582(1991).  
 SQ SEQUENCE 12 AA; 1335 MW; CC9702EC3C233DC2 CRC64;

Query Match 24.3%; Score 18; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 GEY 13  
 Db 1 GEY 3  
 I:|

RESULT 6

Q9RFZ4 ID Q9RFZ4 PRELIMINARY; PRT; 13 AA.

AC Q9RFZ4;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 1.5 KDA PROTEIN (FRAGMENT).  
 OS Mycoplasma mycoides subsp. capri.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Entomoplasmataceae.  
 OX NCBI\_TaxID=40477;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PG3;  
 RA Thiaucourt F., Lorenzon S., David A.;  
 RT "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing

RT of a putative membrane protein gene."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 KW EMBL; AF162998; AAF15253.1; -  
 DR Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 13 AA; 1505 MW; 0B79431F5635573B CRC64;

Query Match 24.3%; Score 18; DB 2; Length 13;  
 Best Local Similarity 57.18; Pred. No. 1e+04; Indels 0; Gaps 0;  
 Matches 4; Conservative 1; Mismatches 2;

QY 4 GIRYVQK 10  
 II I I  
 Db 3 GIPYLIK 9

RESULT 7  
 Q9BF92 PRELIMINARY; PRT; 8 AA.  
 ID Q9BF92  
 AC Q9BF92  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).  
 GN CREM.  
 OS Tursiops truncatus (Atlantic bottle-nosed dolphin).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;  
 OC Tursiops.  
 OX NCBI\_TaxID=9739;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082082; PubMed=11214319;  
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 RA O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals."  
 RL Nature 409:614-618(2001).  
 DR EMBL; AY011670; AAG47581.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 997 MW; DF1DD3240EAB572A CRC64;

Query Match 23.0%; Score 17; DB 6; Length 8;  
 Best Local Similarity 50.08; Pred. No. 4.7e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 YVQKE 12  
 I I I  
 Db 3 YCRKAE 8

RESULT 8  
 Q9XT05 PRELIMINARY; PRT; 9 AA.  
 ID Q9XT05  
 AC Q9XT05  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE ACETYL-COA ACETYLTRANSFERASE 2 (FRAGMENT).  
 GN ACAT2.  
 OS Macropus rufogriseus (Red-necked wallaby).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
 OX NCBI\_TaxID=9320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99282512; PubMed=10353914;  
 RA Shintani S., O'Huigin C., Toyosawa S., Michalova V., Klein J.;  
 RT "Origin of gene overlap: the case of TCPI and ACAT2."  
 RL Genetics 152:743-754(1999).  
 DR EMBL; AF143499; AAD34976.1; -  
 KW Transferase.

FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 934 MW; 12FF16C3CEB69DD0 CRC64;  
 Query Match 23.0%; Score 17; DB 6; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 4.7e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VQKG 11  
 II I I  
 Db 6 VQKG 9

RESULT 9  
 Q9UMR0 PRELIMINARY; PRT; 12 AA.  
 ID Q9UMR0  
 AC Q9UMR0  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE FRUCTOSE-1,6-BISPHOSPHATASE (EC 3.1.3.11) (FRAGMENT).  
 GN FBP2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tillmann H., Eschrich K.;  
 RT "Structure of human FBP2 gene."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ238483; CAB53359.1; -  
 DR HSSP; P00636; 1FRP.  
 KW Hydrolase.  
 FT NON\_TER 1  
 FT NON\_TER 12  
 SQ SEQUENCE 12 AA; 1382 MW; 4CEB259E57386403 CRC64;

Query Match 23.0%; Score 17; DB 4; Length 12;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+04;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 9 QKGEY 13  
 :: I I  
 Db 5 KRGKY 9

RESULT 10  
 Q9IF00 PRELIMINARY; PRT; 12 AA.  
 ID Q9IF00  
 AC Q9IF00  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE E1A NUCLEOPROTEIN (FRAGMENT).  
 OS Human adenovirus type 5.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=28285;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hayashi S., Gillam I.C., McDonald T., Way D., Harris T.,  
 RA Sedgwick E.G.;  
 RT "E1A DNA of group C adenovirus integrates into human chromosomes."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF288220; AAF91488.1; -  
 KW Nucleocapsid.  
 FT NON\_TER 12  
 SQ SEQUENCE 12 AA; 1337 MW; 06D66D0D4171EEB0 CRC64;

Query Match 23.0%; Score 17; DB 12; Length 12;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+04;



Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRDGI 5  
I I:  
Db 6 CHGGV 10

## RESULT 11

Q9IEZ9 PRELIMINARY; PRT; 12 AA.  
AC Q9IEZ9;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE E1A NUCLEOPROTEIN (FRAGMENT).  
OS Human adenovirus type 5.  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_TaxID=28285;  
RN [1]

RP SEQUENCE FROM N.A.

RA Hayashi S., Gillam I.C., McDonald T., Way D., Harris T.,  
RA Sedgwick E.G.;  
RT "E1A DNA of group C adenovirus integrates into human chromosomes.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF288641; AAF91494.1; -;  
KW Nucleocapsid.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1337 MW; 06D66D0D4171EBB0 CRC64;

Query Match 23.0%; Score 17; DB 12; Length 12;

Best Local Similarity 40.0%; Pred. No. 1.3e+04;

Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRDGI 5  
I I:  
Db 6 CHGGV 10

## RESULT 12

Q9TU76 PRELIMINARY; PRT; 13 AA.  
AC Q9TU76;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE INTERLEUKIN 1 ALPHA (FRAGMENT).  
GN IL1A.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]

RP SEQUENCE FROM N.A.

RA Maddox J.F., Hawken R.J., Matthew P., Davies K.P.;  
RT "Single strand conformational polymorphisms (SSCPs) in the ovine IL1A  
RT and IL6 genes.";  
RL Anim. Genet. 30:317-318(1999).  
DR EMBL; AF117652; AAD25050.1; -;  
FT NON\_TER 1 1  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1520 MW; C26BG6198305DB5D CRC64;

Query Match 23.0%; Score 17; DB 6; Length 13;

Best Local Similarity 37.5%; Pred. No. 1.5e+04;

Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DGIRYVQK 10  
I I I:  
Db 3 DMAYVYSE 10

## RESULT 13

Q9MQK6 PRELIMINARY; PRT; 13 AA.  
AC Q9MQK6;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE CYTOCHROME OXIDASE SUBUNIT 2 (FRAGMENT).  
GN COI1.

OS Rupicapra rupicapra (Chamois).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Rupicapra.

OX NCBI\_TaxID=34869;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PERIPHERAL BLOOD;

RX MEDLINE=20104876; PubMed=10641890;

RA Saullie E., Di Pasquale S., Tartaglia M.;

RT "Rapid communication: nucleotide sequence of chamois, alpine ibex, and  
RL red deer tRNA(Lys) and ATPase8 mitochondrial genes.";

RL J. Anim. Sci. 77:3398-3399(1999).

DR EMBL; AF104681; AAF43477.1; -;

KW Mitochondrion.

FT NON\_TER 1 1

SQ SEQUENCE 13 AA; 1574 MW; 6226A044FCB3733B CRC64;

Query Match 23.0%; Score 17; DB 8; Length 13;

Best Local Similarity 33.3%; Pred. No. 1.5e+04;

Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 IRYVQK 10  
I I I:  
Db 2 LKYFEK 7

## RESULT 14

Q9CU06 PRELIMINARY; PRT; 13 AA.  
AC Q9CU06;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE ZONA PELLUCIDA BINDING PROTEIN (FRAGMENT).  
GN ZPBP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=TESTIS;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Mikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK018876; BAB31470.1; -.  
 DR MGD; MGI:1855701; Zppp.  
 FT NON\_TER 1  
 SQ SEQUENCE 13 AA; 1404 MW; D6A4220999576B42 CRC64;

Query Match 23.0%; Score 17; DB 11; Length 13;  
 Best Local Similarity 30.0%; Pred. No. 1.5e+04;  
 Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRDGIRYVOK 10  
 | : : | |  
 Db 1 CANSUVYGAK 10

## RESULT 15

Q9BFC3 PRELIMINARY; PRT; 8 AA.  
 AC Q9BFC3;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).  
 GN CREM.  
 OS Didelphis marsupialis virginiana (North American opossum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OX NCBI\_TaxID=9267;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082082; PubMed=11214319;  
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 RA O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals."  
 RL Nature 409:614-618(2001).  
 DR EMBL; AY011620; AAG47535.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 978 MW; DF1DD331EEAB572A CRC64;

Query Match 21.6%; Score 16; DB 6; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 YVQKGE 12  
 | | | |  
 Db 3 YCHKAE 8

Search completed: February 21, 2002, 16:59:51  
 Job time: 179 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:53:32 ; Search time 23.68 Seconds  
(without alignments)  
40.665 Million cell updates/sec

Title: US-08-753-851-9  
Perfect score: 74  
Sequence: 1 CRDGIRYVQKGEY 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 149706

Minimum DB seq length: 0  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	41.9	9	15 AAY38367	PAP-derived HLA-bi
2	31	41.9	9	20 AAY45949	Immunogenic peptid
3	30	40.5	13	20 AAY48805	Membrane dipeptida
4	28	37.8	12	16 AAR81762	Vpr/Vpx motif-deri
5	27	36.5	13	21 AAB11009	Immunogenic peptid
6	26	35.1	7	16 AAR79083	Alphas/betal integ
7	26	35.1	10	21 AAY81289	Linker g3-2 peptid
8	26	35.1	13	20 AAY48795	Membrane dipeptida
9	25	33.8	8	11 AAR07967	Tryptic fragment T
10	25	33.8	10	20 AAW95554	Aspergillus tubige
11	25	33.8	10	22 AAG94423	Human complementar

12	25	33.8	10	22	AAG96391	Human complementar
13	25	33.8	12	18	AAW13796	Hepatitis C virus
14	24	32.4	9	8	AAP70356	Sequence of new bo
15	24	32.4	10	16	AAR72230	Epitope of rat vas
16	24	32.4	10	22	AAG95958	Human complementar
17	24	32.4	11	12	AAW14572	Blood platelet agg
18	24	32.4	11	17	AAW02116	Blotylated pv pe
19	24	32.4	11	21	AAY58705	Antiangiogenic pep
20	24	32.4	11	22	AAB80740	Semenogelin II hK2
21	24	32.4	12	16	AAR81782	SIV strain mac239
22	24	32.4	12	16	AAR66971	Human IL-4 mutant
23	24	32.4	12	18	AAW27417	CDR3 from heavy ch
24	24	32.4	12	21	AAY93853	Reactive peptide w
25	24	32.4	13	19	AAW46491	Human growth arres
26	23	31.1	7	21	AAY64089	Desmocollin cell a
27	23	31.1	8	16	AAW21280	Hydroxymethylgluta
28	23	31.1	8	16	AAR67903	Anti-2-phenylloxazo
29	23	31.1	8	20	AAW97448	Peptide that is us
30	23	31.1	9	15	AAY38357	PAP-derived HLA-bi
31	23	31.1	9	15	AAW48424	Histogranin deriva
32	23	31.1	9	20	AAY45938	Immunogenic peptid
33	23	31.1	9	21	AAW32283	HLA A1 binding TAD
34	23	31.1	9	22	AAE00872	Mutant leukocyte I
35	23	31.1	10	13	AAW24683	Immunomodulatory p
36	23	31.1	10	15	AAY38363	PAP-derived HLA-bi
37	23	31.1	10	15	AAW48418	Histogranin deriva
38	23	31.1	10	18	AAW10903	Polyclonal anti-HB
39	23	31.1	10	20	AAY45945	Immunogenic peptid
40	23	31.1	10	20	AAY46869	Immunogenic peptid
41	23	31.1	10	20	AAW78436	Human ADAMTS-1 pep
42	23	31.1	10	22	AAG83364	Arabidopsis thalia
43	23	31.1	10	22	AAG88258	Saccharomyces cere
44	23	31.1	10	22	AAG88259	Saccharomyces cere
45	23	31.1	11	15	AAW48423	Histogranin deriva

## ALIGNMENTS

RESULT 1  
AAY38367  
ID AAY38367 standard; Peptide; 9 AA.  
AC AAY38367;  
XX  
DT 29-SEP-1999 (first entry)  
XX  
DE PAP-derived HLA-binding peptide.  
XX  
KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;  
KW MHC; major histocompatibility complex; viral infection; anticancer;  
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.  
XX  
OS Homo sapiens.  
XX  
PN WO9403205-A1.  
XX  
PD 17-FEB-1994.  
XX  
PF 06-AUG-1993; 93WO-US07421.  
XX  
PR 05-MAR-1993; 93US-0027746.  
PR 07-AUG-1992; 92US-0926666.  
(CYTE-) CYTEL CORP.  
Celis E, Grey HM, Kubo RT, Sette A;  
WPI; 1994-065403/08.  
XX  
PT Peptide which specifically binds selected MHC allele - used to induce an immune response for treatment or prevention of viral infection or cancer, or for diagnosis

XX PS Disclosure; Page 114; 150pp; English.

XX CC The sequence is a specific example of a group of new immunogenic

CC peptides having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding

CC motif. For example, the peptides having an HLA-A3.2 binding motif

CC each have 9-10 residues and contain, from the N-terminus to the

CC C-terminus, (a) a first conserved residue selected from L, M, I,

CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of

CC K, R, Y, H or F, where the first and second conserved residues are

CC separated by 6-7 residues. The peptides are capable of binding

CC selected MHC molecules and inducing an immune response. They can be

CC used to treat and/or prevent viral infection and cancer, e.g. prostate

CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce

CC antibodies for use as diagnostic or therapeutic agents. The peptides

CC can also be used as diagnostic agents.

XX SQ Sequence 9 AA;

Query Match 41.9%; Score 31; DB 15; Length 9;

Best Local Similarity 71.4%; Pred. No. 4.3e+05;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 YVQKGEY 13

DB 1 :||||

2 yfekgey 8

RESULT 2

AAAY45949

ID AAY45949 standard; Peptide; 9 AA.

XX AC AAY45949;

XX DT 01-DEC-1999 (first entry)

XX DE Immunogenic peptide having a human leukocyte antigen binding motif #560.

XX KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;

XX KW immune response; T cell activation; major histocompatibility complex;

XX KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;

XX KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;

XX KW vaccine; immunisation.

XX OS Synthetic.

OS Homo sapiens.

PN WO9945954-A1.

XX PD 16-SEP-1999.

XX PF 13-MAR-1998; 98WO-US05039.

XX PR 13-MAR-1998; 98WO-US05039.

XX PA (EPTM-) EPIMUNE INC.

PI Sette A, Kubo RT, Sidney J, Cells E, Grey HM, Southwood S;

PI WPI; 1999-551214/46.

DR New immunogenic peptides with HLA binding motif, useful in treatment

XX and diagnosis of cancers and viral diseases -

PS Claim 1; Page 48; 150pp; English.

XX CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides

CC having a human major histocompatibility complex (MHC) Class I (also

CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic

CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes

CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell

CC response against the antigen from which the peptide is derived.

CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are

CC normally induced by an antigen in the form of a peptide fragment bound

CC to a HLA molecule, rather than the intact foreign antigen itself, and

CC are particularly important in tumour rejection and in fighting viral

CC infections. The peptides are therefore useful therapeutically to treat

CC or prevent viral infections and cancers in mammals (especially humans)

CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.

CC They can be administered as vaccines to elicit an immune response in

CC individuals susceptible or otherwise at risk of viral infection or

CC cancer, or used to treat chronic or acute conditions. They are also

CC useful diagnostically, and can be used to induce a cytotoxic T cell

CC response, by contacting a cytotoxic T cell with the peptide e.g. to

CC produce CTLs ex vivo for infusion back into a patient. The

CC polynucleotides encoding the immunogenic peptides are also useful

CC therapeutically and for immunisation as above.

XX SQ Sequence 9 AA;

Query Match 41.9%; Score 31; DB 20; Length 9;

Best Local Similarity 71.4%; Pred. No. 4.3e+05;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 YVQKGEY 13

DB 1 :||||

2 yfekgey 8

RESULT 3

AAAY48805

ID AAY48805 standard; Peptide; 13 AA.

XX AC AAY48805;

XX DT 10-DEC-1999 (first entry)

XX DE Membrane dipeptidase-binding skin homing peptide #21.

XX KW Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;

XX KW prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;

XX KW membrane dipeptidase.

XX OS Synthetic.

OS Homo sapiens.

PN WO9946284-A2.

XX PD 16-SEP-1999.

XX PF 10-MAR-1999; 99WO-US05284.

XX PR 13-MAR-1998; 98US-0042107.

XX PR 26-FEB-1999; 99US-0042107.

XX PA (BURN-) BURNHAM INST.

XX PI Rajotte D, Pasqualini R, Ruoslahti EI;

XX DR WPI; 1999-571717/48.

XX PT New peptides which selectively home to organs or tissues, used for,

PT e.g. identifying target ligands and for therapy of pathological

XX conditions -

PS Claim 19; Page 148; 193pp; English.

XX CC The present invention describes peptides that selectively home to a

CC tissue or organ. The peptides can be used for identifying an organ

CC or tissue, for identifying a target molecule expressed by an organ or

CC tissue or for treating an organ or tissue pathology, where the organ or

CC tissue is selected from prostate, lung, skin, retina, pancreas, gut,

CC ovary, adrenal gland, liver, and lymph node. The peptide bind to the

CC membrane dipeptidase (MDP). AAY48618 to AAY49066 represent sequences

CC which are used in the exemplification of the present invention.

XX Sequence 13 AA;

Query Match 40.5%; Score 30; DB 20; Length 13;

Best Local Similarity 50.0%; Pred. No. 48;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CRDGRYV 8

Db 5 cadgrcfi 12

RESULT 4

ID AAR81762 standard; peptide; 12 AA.

XX AC AAR81762;

DT 10-MAY-1996 (first entry)

DE Vpr/Vpx motif-derived antiproliferative peptide 6.

XX Motif: Vpr; Vpx; human immunodeficiency virus; HIV; virion-associated;  
KW infection; penetration; uncoating; conserved; primate; lentivirus;  
KW antagonist; anti-proliferative; eukaryotic cell; cancer; leukaemia;  
KW psoriasis.

XX OS Synthetic.

XX PN WO9526361-A1.

XX PD 05-OCT-1995.

XX PF 24-MAR-1995; 95WO-AU00169.

XX PR 03-FEB-1995; 95AU-0000902.

XX PS 25-MAR-1994; 94AU-0004697.

XX PA (BIOM-) BIOMOLECULAR RES INST LTD.

XX PI Arunagiri C, Azad AA, Macreadie IG;

XX WPI; 1995-351293/45.

XX Vpr and Vpx proteins of HIV, and derived peptide(s) - useful as  
PT antagonists for the treatment of HIV, cell proliferative diseases  
PT and diseases caused by pathogens

PS Example 6; Page 15; 69pp; English.

XX Peptides AAR81757-62 are synthesised to contain the human  
CC immunodeficiency virus (HIV) Vpr peptide consensus motif (see  
CC AAR81732-3). The peptides, which are synthesised to contain 0, 1 or 2  
CC consensus motifs, are used to osmosensitise yeast cells based on the  
CC antiproliferative activity of the Vpr protein, esp. the C-terminal motif.  
CC Peptides derived from the Vpr/Vpx C-terminal motifs can be used in the  
CC treatment of diseases mediated by cell proliferation such as cancer,  
CC leukaemia or psoriasis.

XX Sequence 12 AA;

Query Match 37.8%; Score 28; DB 16; Length 12;

Best Local Similarity 50.0%; Pred. No. 1e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 RDGIRYVQKG 11

Db 3 ragtryfrfg 12

RESULT 5

AAB11009

ID AAB11009 standard; peptide; 13 AA.

XX AC AAB11009;

DT 08-FEB-2001 (first entry)

DE Immunogenic peptide #1 for raising human C3a antigen.

XX Immunogenic; antigen; C3a; C5a; IL10; IL1-alpha; IL1-beta; IL6;  
KW TNF-alpha; immunoadsorber; treatment; sepsis; complement factor;  
KW anti-inflammatory; antiseptis; plasmapheresis; septic shock.

XX OS Homo sapiens.

XX PN DE19913707-A1.

XX PD 05-OCT-2000.

XX PF 26-MAR-1999; 99DE-1013707.

XX PR 26-MAR-1999; 99DE-1013707.

XX PA (PRIV-) PRIVATES INST BIOSERV GMBH.

XX PI Heinrich H, Hahn H, Meyer U, Kruschke P, Wagner H;

XX WPI; 2000-648176/63.

XX Immunoadsorber for treating sepsis, comprises polymeric support  
PT carrying antibodies specific for complement factors and  
PT lipopolysaccharide

XX Claim 6; Page 4; 8pp; German.

XX This invention describes a novel immunoadsorber (A) for treating sepsis  
CC which comprises a carrier of organic or synthetic polymer having bound to  
CC it poly- or monoclonal antibodies (Ab1) against the complement factors  
CC C3a and/or C5a and against lipopolysaccharide (LPS). Optionally other  
CC antibodies (Ab2) directed against other mediators of sepsis are also  
CC attached to the carrier. The products of the invention have  
CC anti-inflammatory and antiseptic activity. (A) are used for treatment,  
CC particularly by plasmapheresis, of sepsis and septic shock. The  
CC combination of antibodies used can be tailored to the requirements of  
CC individual patients.

XX Sequence 13 AA;

Query Match 36.5%; Score 27; DB 21; Length 13;

Best Local Similarity 66.7%; Pred. No. 1.7e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CRDGIR 6

Db 3 cedgmr 8

RESULT 6

AAR79083

ID AAR79083 standard; peptide; 7 AA.

XX AC AAR79083;

DT 24-JAN-1996 (first entry)

DE Alpha5/beta1 integrin binding peptide #13.

XX High affinity; integrin binding peptide; alpha5/beta1; alphav/beta5;  
KW alphav/beta3; RGD; stable configuration; wound healing;  
KW osteoclast attachment; bone; angiogenesis; metastasis; tumour;  
KW smooth muscle cell migration.

```

XX OS Synthetic.
XX PN WO9514714-A1.
XX PD 01-JUN-1995.
XX PF 22-NOV-1994; 94WO-US13542.
XX PR 04-AUG-1994; 94US-0286861.
XX PR 24-NOV-1993; 93US-0158001.
XX PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX PI Koivunen E, Ruoslahti E;
XX DR WPI; 1995-206899/27.
XX PT High affinity integrin binding peptides - can be used to attach
XX PT cells to a substrate, inhibit the attachment of osteoclasts to bone,
XX PT promote wound healing, inhibit angiogenesis, metastasis of tumours
XX PT and migration of smooth muscle cells
XX PS Example 9; Page 34; 86pp; English.
XX CC The sequences given in AAR76185-200 and AAR79073-94 are high affinity
XX CC integrin binding peptides which bind to various integrins. Peptides
XX CC which bind to alpha5/beta1 integrins contain the motifs given in
XX CC AAR76185-86 and peptides which bind to alpha5/beta5 and alpha5/beta3
XX CC integrins contain the motif given in AAR76187. Alpha5/beta5 integrins
XX CC are also bound by RGD containing peptides. These peptides assume a
XX CC conformationally stabilised configuration which is due to the
XX CC formation of a disulphide bond, a peptide bond or a lactam bond.
XX CC These peptides may be used for isolating the complementary integrin
XX CC from a sample mixture by contacting them under ionic conditions to
XX CC allow binding of the integrin to the peptide and then separating the
XX CC integrin from the peptide. They can be used for attaching cells to
XX CC a substrate, by binding them to the substrate with the cell. The
XX CC peptides promote wound healing when applied locally and inhibit the
XX CC attachment of osteoclasts to bone. They inhibit angiogenesis,
XX CC metastasis of tumours and migration of smooth muscle cells.
XX CC Sequence 7 AA;
XX QY 1 CRDG 4
XX DB 1 crdg 4
XX RESULT 7
XX AAY81289
XX ID AAY81289 standard; peptide; 10 AA.
XX AC AAY81289;
XX DT 05-JUN-2000 (first entry)
XX DE Linker g3-2 peptide, SEQ ID NO:136.
XX KW Phase display; bacteriophage M13; fusion protein; major coat protein;
XX KW protein VIII; phagemid vector; electroporation; combinatorial library;
XX OS protein III.
XX OS Synthetic.
XX PN WO200006717-A2.
XX PD 10-FEB-2000.

```

```

XX 22-JUL-1999; 99WO-US16596.
XX PF 27-JUL-1998; 98US-0094291.
XX PR 08-OCT-1998; 98US-0103514.
XX PR 10-MAY-1999; 99US-0133296.
XX PR 19-MAY-1999; 99US-0134870.
XX PA (GETH ) GENENTECH INC.
XX PI Sidhu SS, Weiss GA, Wells JA;
XX DR WPI; 2000-183122/16.
XX PT Fusion proteins comprising a heterologous protein and a viral variant
XX PT major coat protein useful in phage display systems for improving
XX PT transformation efficiency -
XX PS Example 26; Page 85; 118pp; English.
XX CC The invention relates to novel fusion proteins comprising a heterologous
XX CC polypeptide fused to a variant (non-wild type) bacteriophage to
XX CC major coat protein (protein VIII). The invention also relates to
XX CC replicable expression vectors which contain a gene encoding the fusion
XX CC protein; host cells containing the expression vectors; phages which
XX CC display the fusion protein on their surface; phage libraries displaying
XX CC a plurality of different fusion proteins on viral surfaces; and methods
XX CC of using these compositions. The fusion proteins the invention are well
XX CC tolerated in phage display systems. Variants of the major coat proteins
XX CC can be used to alter the number of fusion proteins incorporated into a
XX CC virus particle. Hyper-functional variants can be used to increase the
XX CC number of fusion proteins incorporated into a virus particle.
XX CC Conversely, hypo-functional variants can be used to decrease fusion
XX CC protein incorporation. This is useful for tailoring the incorporation of
XX CC fusion proteins into virus particles to achieve a desired level of
XX CC valency. The variant replicable plasmid/phagemid vectors are useful for
XX CC producing polypeptides of interest. The methods are useful for improving
XX CC the transformation of cells by highly purifying DNA. The present
XX CC invention uses affinity DNA purification to reduce ionic impurities and
XX CC thus reduce the conductance associated with a unit mass of DNA. This is
XX CC an advantageous in electroporation methods for increasing the
XX CC concentration of DNA present. The increase in DNA entering the host cell
XX CC provides a greater number of transformants per electroporation and allows
XX CC one to prepare larger combinatorial libraries which overcomes the prior
XX CC art problem of small library size using recombinant DNA. Sequences
XX CC AAY81279-Y81283 and AAY81288-Y81290 represent linkers selected for
XX CC display of a peptide fused to the C-terminal end of protein VIII
XX CC (AAY81279-Y81283) or protein III (AAY81288-Y81290).
XX QY Sequence 10 AA;
XX DB 5 IRYVORKE 12
XX DB 2 irvlqkgk 9
XX Query Match 35.1%; Score 26; DB 21; Length 10;
XX Best Local Similarity 62.5%; Pred. No. 1.9e+02;
XX Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX QY 5 IRYVORKE 12
XX DB 2 irvlqkgk 9
XX RESULT 8
XX AAY48795
XX ID AAY48795 standard; peptide; 13 AA.
XX AC AAY48795;
XX DT 10-DEC-1999 (first entry)
XX DE Membrane dipeptidase-binding skin homing peptide #11.
XX KW Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;
XX KW prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;

```

KW membrane dipeptidase.

XX Synthetic.  
OS Homo sapiens.

XX WO9946284-A2.

XX PD 16-SEP-1999.

XX PF 10-MAR-1999; 99WO-US05284.

XX PR 13-MAR-1998; 98US-0042107.

XX PR 26-FEB-1999; 99US-0042107.

XX (BURN-) BURNHAM INST.

PI Rajotte D, Pasqualini R, Ruoslahti EI;

XX WPI; 1999-571717/48.

XX New peptides which selectively home to organs or tissues, used for,  
PT e.g. identifying target ligands and for therapy of pathological  
PT conditions

XX Example 6; Page 148; 193pp; English.

XX The present invention describes peptides that selectively home to a  
CC tissue or organ. The peptides can be used for identifying an organ  
CC or tissue, for identifying a target molecule expressed by an organ  
CC tissue or for treating an organ or tissue pathology, where the organ or  
CC tissue is selected from prostate, lung, skin, retina, pancreas, gut,  
CC ovary, adrenal gland, liver, and lymph node. The peptide bind to the  
CC membrane dipeptidase (MDP). AAY48618 to AAY49066 represent sequences  
CC which are used in the exemplification of the present invention.

XX Sequence 13 AA;

Query Match 35.1%; Score 26; DB 20; Length 13;  
Best Local Similarity 50.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRDGRYV 8

Db 5 cadgcrl 12

RESULT 9

AAR07967  
ID AAR07967 standard; protein; 8 AA.

XX AC AAR07967;

XX DT 14-JAN-1991 (first entry)

XX Tryptic fragment T17 (b) of rat phospholipase A2 inhibitor.

XX Human N-lipocortin; placenta; inflammation reduction; arthritis;  
KW rat phospholipase A2 inhibitor; tryptic fragment T17 (b).

XX OS Homo sapiens.

XX PN US4950646-A.

XX PD 21-AUG-1990.

XX PF 10-JAN-1986; 86US-0929199.

XX PR 10-JAN-1986; 86US-0929199.

XX PR 05-SEP-1985; 85US-0772892.

XX PR 14-AUG-1985; 85US-0765877.

XX PR 15-MAR-1985; 85US-0712376.

XX PR 10-JAN-1985; 85US-0690146.

XX PA

(BIOJ ) BIOGEN NV.

XX Wallner BP, Pepinsky RB, Garwin JL, Schindler DG, Huang KS;

XX WPI; 1990-274549/36.

XX Pure fragment of human lipocortin - useful for reducing  
PT inflammation or for treating arthritis, etc.

XX PS Disclosure; Fig 2; 51pp; English.

XX Rat phospholipase A2 inhibitor protein was isolated from the  
CC extracellular supernatant of rat peritoneal exudate cells.

CC For tryptic fragment T17 (a) see AAR07929.

CC The protein was isolated and sequenced to produce oligonucleotide  
CC probes in order to identify human lipocortin and N-lipocortin.

CC See also AAQ05805-25, AAQ06581, AAR07926-37 and AAR07956-66.

XX Sequence 8 AA;

Query Match 33.8%; Score 25; DB 11; Length 8;

Best Local Similarity 66.7%; Pred. No. 4.3e+05;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 YVOKGE 12

Db 2 fvqkgq 7

RESULT 10

AAW95554

ID AAW95554 standard; Peptide; 10 AA.

XX AC AAW95554;

XX DT 10-MAY-1999 (first entry)

XX Aspergillus tubigenensis xylanase N-terminal peptide.

XX Endoxylanase A; XlnA; xylanase; brewing; baking; pulp; paper;  
KW lignin removal; delignification.

XX OS Aspergillus tubigenensis.

XX PN EP892065-A1.

XX PD 20-JAN-1999.

XX PF 24-JUL-1991; 91EP-0201944.

XX PR 24-JUL-1990; 90EP-0202020.

XX (KONN ) GIST-BROCADES NV.

XX WPI; 1999-083582/08.

XX New fungal xylanase gene - useful in the preparation of bread,  
PT animal foods, and in paper products

XX Example 1.2; Fig 3; 55pp; English.

XX This is an N-terminal peptide of a 25 kDa xylanase isolated from  
CC the culture filtrate of Aspergillus tubigenensis. Probes (see  
CC AAX00982-87) based on residues 4-10 of the peptide were used to  
CC screen a DNA library prepared from Aspergillus niger DS16813 (now  
CC renamed Aspergillus tubigenensis CBS 323.90) to isolate a xlnA gene  
CC (see AAX00973) encoding a new endoxylanase A (see AAW95553). The  
CC endoxylanase shows optimum activity at pH 3.5-5.5, making it  
CC useful for industrial processes requiring lower pH, such as  
CC lagering of beer, and can also be used in baking, preparation of  
CC animal feedstuff, lignin removal from kraft mill pulp, and in the

CC preparation of paper products (all claimed).

XX Sequence 10 AA;

Query Match 33.8%; Score 25; DB 20; Length 10;

Best Local Similarity 83.3%; Pred. No. 2.9e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GIRYVQ 9

Db 11111

2 ginyvq 7

RESULT 11

AAG94423

ID AAG94423 standard; Peptide; 10 AA.

XX AC AAG94423;

DT 18-SEP-2001 (first entry)

DE Human complementary peptide, SEQ ID NO: 617.

KW Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

XX WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB04776.

PR 13-DEC-1999; 99GB-0029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

WPI; 2001-408419/43.

A set of peptide ligands consisting of specific complementary peptides to proteins encoded by genes of the human genome, useful in an assay for screening and identifying of one or more novel peptides which are drug candidates or pro-drugs -

Example 4; Page 128; 646pp; English.

The invention relates to a set of complementary peptide ligands

generated from the human genome. The complementary peptides

interact with their relevant target proteins encoded in the human

genome. They can be used as reagents in drug discovery and as lead

ligands to facilitate drug design and development. The present

sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match

Best Local Similarity 33.8%; Score 25; DB 22; Length 10;

Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GIRYVQKGEY 13

Db 11111

1 ginyctsgay 10

RESULT 12

AAG96391

ID AAG96391 standard; Peptide; 10 AA.

XX AC AAG96391;

XX 18-SEP-2001 (first entry)

DE Human complementary peptide, SEQ ID NO: 2585.

KW Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

XX WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB04776.

PR 13-DEC-1999; 99GB-0029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

WPI; 2001-408419/43.

A set of peptide ligands consisting of specific complementary peptides to proteins encoded by genes of the human genome, useful in an assay for screening and identifying of one or more novel peptides which are drug candidates or pro-drugs -

Example 4; Page 412; 646pp; English.

The invention relates to a set of complementary peptide ligands

generated from the human genome. The complementary peptides

interact with their relevant target proteins encoded in the human

genome. They can be used as reagents in drug discovery and as lead

ligands to facilitate drug design and development. The present

sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match

Best Local Similarity 33.8%; Score 25; DB 22; Length 10;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 QKGEY 13

Db 4 ekgey 8

RESULT 13

AAG13796

ID AAG13796 standard; Peptide; 12 AA.

XX AC AAG13796;

DT 10-NOV-1997 (first entry)

Hepatitis C virus NS4A/NS4B junction region derivative peptide.

NS3; protease; HCV; inhibitor; antiviral; virucide.

OS Synthetic.

XX Key Modified-site 1

Location/Qualifiers

/label= OTHER

/note= "OTHER = Ac-Glu"

XX WO9708304-A2.

XX 06-MAR-1997.

XX 20-AUG-1996; 96WO-IT00163.



XX 22-AUG-1995; 95IT-RM00573.  
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
 XX Bianchi E, De Francesco R, Narjes F, Pessi A, Steinkuehler C;  
 PI Talliani M, Tomei L, Urbani A;  
 XX WPI; 1997-179261/16.  
 XX New isolated hepatitis C virus NS3 polypeptide(s) - useful in  
 PT high-throughput assays for detecting inhibitors which can be used to  
 PT inhibit virus replication  
 XX Example 3; Page 18; 70pp; English.  
 PS A synthetic peptide (AAW13795) corresponds to an N-terminally deleted  
 XX version of the hepatitis C virus (HCV) polyprotein NS4A/4B junction  
 CC (see also AAW13771). Unlike other related peptides (AAW13771-86) and  
 CC desipeptides (AAW13787-91), this peptide is not suitable as a  
 CC substrate for assays of HCV NS3 protease activity.  
 XX Sequence 12 AA;  
 SQ  
 Query Match 33.8%; Score 25; DB 18; Length 12;  
 Best Local Similarity 27.3%; Pred. No. 3.6e+02;  
 Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CRDGIRYVQKG 11  
 Db | : : : : :  
 2 cashipyiegg 12  
 RESULT 14  
 AAP70356  
 ID AAP70356 standard; peptide; 9 AA.  
 XX AC AAP70356;  
 XX 26-APR-1991 (first entry)  
 DT  
 XX Sequence of new bone morphogenetic agent.  
 DE  
 XX Bone morphogenetic protein; osteoinduction; bone growth;  
 KW cartilage growth.  
 XX EP212474-A.  
 PN  
 XX 04-MAR-1987.  
 PD  
 XX 07-AUG-1986; 86EP-0110935.  
 PF  
 XX 07-AUG-1985; 85US-0763479.  
 PR 04-AUG-1980; 80US-0174906.  
 PR 05-MAY-1981; 81US-0260726.  
 PR 16-AUG-1983; 83US-0523606.  
 XX (RECC ) UNIV OF CALIFORNIA.  
 PA  
 XX Urst MR;  
 PI  
 XX WPI; 1987-058071/09.  
 DR  
 XX New bone morphogenetic agent comprising osteo-inductive peptide -  
 PT is obtd. by recombinant DNA procedures etc. giving larger amts.  
 PT for use in humans and animals  
 XX Claim 5; p14; 17pp; English.  
 PS A bone morphogenetic agent comprising a peptide displaying  
 XX osteoinductive activity similar to bone morphogenetic protein (BMP)  
 CC is new. The peptide may be obtained from any source, including

CC natural, synthetic or recombinant DNA sources. The new agent causes  
 CC bone tissues and cartilage to form and/or grow in humans and animals.  
 XX Sequence 9 AA;  
 SQ  
 Query Match 32.4%; Score 24; DB 8; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CRDGIR 6  
 Db | : : : :  
 2 cadgyr 7  
 RESULT 15  
 AAR72230  
 ID AAR72230 standard; Peptide; 10 AA.  
 XX AC AAR72230;  
 XX 01-DEC-1995 (first entry)  
 DT  
 XX Epitope of rat vascular smooth muscle AT1 receptor (residues 8-17).  
 DE  
 XX Monoclonal antibody; AT1 receptor; subtype; Angiotensin II receptor;  
 KW rat vascular smooth muscle; conserved; mammalian AT1 receptor;  
 KW detection; cancer diagnosis; sperm motility; contraception.  
 XX OS Synthetic.  
 XX WO9509186-A.  
 PN  
 XX 06-APR-1995.  
 PD  
 XX 27-SEP-1994; 94WO-GB02100.  
 PF  
 XX 27-SEP-1993; 93GB-0019877.  
 PR  
 XX (QUEE-) QUEEN MARY & WESTFIELD COLLEGE.  
 PA  
 XX Barker S, Vinson GP;  
 PI  
 XX WPI; 1995-147395/19.  
 DR  
 XX Monoclonal antibody to AT1 angiotensin receptor - used for  
 PT detection and control of uterine contractions and  
 PT vasoconstriction e.g. for treating hypertension  
 XX Claim 3; Page 13; 24pp; English.  
 PS  
 XX Monoclonal antibodies (MAbs) that bind to the AT1 receptor (a  
 CC subtype of Angiotensin II receptor) are claimed. The Abs  
 CC specifically bind to amino acid residues 8 to 17 of the rat vascular  
 CC smooth muscle AT1 receptor. This sequence is conserved in all  
 CC mammalian AT1 receptors so far cloned. The MAbs can be used for  
 CC detection of AT1 receptors, e.g. in cancer diagnosis. They can also  
 CC be used to study and measure sperm motility and can be used in  
 CC contraception.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 32.4%; Score 24; DB 16; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 DGIRYVQ 9  
 Db | : : : : :  
 2 dgikriq 8

Search completed: February 21, 2002, 16:56:47  
Job time: 195 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 21, 2002, 16:55:12 ; Search time 12.46 seconds  
(without alignments)  
23.479 Million cell updates/sec

Title: US-08-753-851-9

Perfect score: 74

Sequence: 1 CRGIRVQKGEY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 79885

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	41.9	9	3	US-08-159-339A-628
2	30	40.5	13	4	US-09-258-754-186
3	30	40.5	13	4	US-09-042-107-186
4	29	39.2	10	3	US-08-159-339A-928
5	28	37.8	13	2	US-08-867-941-67
6	27	36.5	10	1	US-07-768-2868-17
7	26	35.1	10	1	US-07-842-349-3
8	26	35.1	13	4	US-09-258-754-195
9	26	35.1	13	4	US-09-042-107-195
10	25	33.8	10	2	US-08-867-941-28
11	25	33.8	10	4	US-09-074-658-28
12	25	33.8	12	1	US-08-439-905-7
13	25	33.8	12	3	US-09-100-536-8
14	25	33.8	12	3	US-09-100-537-8
15	25	33.8	12	4	US-09-011-961-17
16	24	32.4	10	3	US-08-624-374-1
17	24	32.4	11	4	US-09-433-598-14
18	24	32.4	11	4	US-09-433-598-15
19	24	32.4	12	2	US-08-591-438-36
20	24	32.4	12	3	US-08-836-561-42
21	24	32.4	13	2	US-08-883-070-5
22	23	31.1	9	2	US-08-637-759B-444
23	23	31.1	9	3	US-08-871-355A-444
24	23	31.1	9	3	US-08-159-339A-618
25	23	31.1	9	4	US-09-393-554-9
26	23	31.1	9	4	US-09-518-046-67
27	23	31.1	10	3	US-08-159-339A-624

28	23	31.1	11	4	US-09-433-598-16
29	23	31.1	11	4	US-09-410-025-10
30	23	31.1	12	1	US-08-250-789A-27
31	23	31.1	13	2	US-08-372-197-1
32	22	29.7	7	4	US-08-946-525-13
33	22	29.7	8	1	US-08-615-181-69
34	22	29.7	8	1	US-08-615-181-75
35	22	29.7	8	4	US-08-160-604-49
36	22	29.7	8	4	US-08-160-604-50
37	22	29.7	8	4	US-08-160-604-51
38	22	29.7	9	1	US-07-944-143C-3
39	22	29.7	9	1	US-08-615-181-56
40	22	29.7	9	2	US-08-986-234-56
41	22	29.7	9	3	US-08-159-339A-944
42	22	29.7	9	3	US-08-676-818-16
43	22	29.7	9	4	US-09-407-549-16
44	22	29.7	9	5	PCT-US93-08214-3
45	22	29.7	10	1	US-07-801-812A-7

#### ALIGNMENTS

RESULT 1  
US-08-159-339A-628  
; Sequence 628, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Celis, Esteban  
; TITLE OF INVENTION: H1A Binding peptides and Their  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-005030US  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 628:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-159-339A-628

Query Match 41.9%; Score 31; DB 3; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.6e+05;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 YVQGEY 13  
| :|||  
Db 2 YFEKGEY 8

## RESULT 2

US-09-258-754-186  
; Sequence 186, Application US/09258754  
; Patent No. 6174687  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Rajotte, Daniel  
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
; TITLE OF INVENTION: Membrane Dipeptidase  
; FILE REFERENCE: P-LJ 3443  
; CURRENT APPLICATION NUMBER: US/09/258,754  
; CURRENT FILING DATE: 1999-02-26  
; EARLIER APPLICATION NUMBER: 09/042,107  
; EARLIER FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 452  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 186  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-258-754-186

Query Match 40.5%; Score 30; DB 4; Length 13;  
Best Local Similarity 50.0%; Pred. No. 18;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRDGIRYV 8  
| :|||  
Db 5 CADGCRFI 12

## RESULT 3

US-09-042-107-186  
; Sequence 186, Application US/09042107  
; Patent No. 6232287  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
; TITLE OF INVENTION: Tissues  
; FILE REFERENCE: P-LJ 2892  
; CURRENT APPLICATION NUMBER: US/09/042,107  
; CURRENT FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 436  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 186  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-042-107-186

Query Match 40.5%; Score 30; DB 4; Length 13;  
Best Local Similarity 50.0%; Pred. No. 18;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRDGIRYV 8  
| :|||  
Db 5 CADGCRFI 12

## RESULT 4

US-08-159-339A-928  
; Sequence 928, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Celis, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-005030US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 928:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-159-339A-928

Query Match 39.2%; Score 29; DB 3; Length 10;  
Best Local Similarity 57.1%; Pred. No. 20;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRDGIRY 7  
| :|||  
Db 4 CRDTLKY 10

## RESULT 5

US-08-867-941-67  
; Sequence 67, Application US/08867941  
; Patent No. 5977337  
; GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M  
APPLICANT: Du, Run-Pan  
APPLICANT: Wang, Quijun  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Klein, Michael H  
TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/867,941  
FILING DATE: 03-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-867-941-67

Query Match 37.8%; Score 28; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 42;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 IRYVKG 12  
::: |||  
Db 2 VQYTRKE 9

RESULT 6  
US-07-768-286B-17  
Sequence 17, Application US/07768286B  
Patent No. 5444153  
GENERAL INFORMATION:  
APPLICANT: GOSS, Neil H.  
APPLICANT: RICHARDSON, Michael A.  
TITLE OF INVENTION: VARIANTS OF PAI-2  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/768,286B  
FILING DATE: 19911011  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00603  
FILING DATE: 20-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/157 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-07-768-286B-17

Query Match 36.5%; Score 27; DB 1; Length 10;  
Best Local Similarity 44.4%; Pred. No. 47;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IRYVQKGEY 13  
::: ||| |  
Db 2 MQQIQKGSY 10

RESULT 7  
US-07-842-349-3  
Sequence 3, Application US/07842349  
Patent No. 5358864  
GENERAL INFORMATION:  
APPLICANT: VAN DEN BROECK, HENRIETTE C.  
APPLICANT: DE GRAAFF, LEENDERT H.  
APPLICANT: HILLE R., JAN D.  
APPLICANT: VAN OOVEN J., ALBERT J.  
APPLICANT: VISSER, JACOB  
APPLICANT: HARDER, ABRAHAM  
TITLE OF INVENTION: CLONING AND EXPRESSION OF XYLANASE-GNEES  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/842,349  
FILING DATE: 19920427  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20031.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids

; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-07-842-349-3

Query Match 35.1%; Score 26; DB 1; Length 10;  
Best Local Similarity 83.3%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 G1RVVQ 9  
| | | |  
DB 2 G1NVVQ 7

RESULT 8  
US-09-258-754-195  
; Sequence 195, Application US/09258754  
; Patent No. 6174687  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Rajotte, Daniel  
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
; FILE REFERENCE: P-LJ 3443  
; CURRENT APPLICATION NUMBER: US/09/258,754  
; CURRENT FILING DATE: 1999-02-26  
; EARLIER APPLICATION NUMBER: 09/042,107  
; EARLIER FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 452  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 195  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-258-754-195

Query Match 35.1%; Score 26; DB 4; Length 13;  
Best Local Similarity 50.0%; Pred. No. 97;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRDGIRVY 8  
| | | |  
DB 5 CADGCRLL 12

RESULT 9  
US-09-042-107-195  
; Sequence 195, Application US/09042107  
; Patent No. 6232287  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
; TITLE OF INVENTION: Tissues  
; FILE REFERENCE: P-LJ 2892  
; CURRENT APPLICATION NUMBER: US/09/042,107  
; CURRENT FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 436  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 195  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-042-107-195

Query Match 35.1%; Score 26; DB 4; Length 13;  
Best Local Similarity 50.0%; Pred. No. 97;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRDGIRVY 8  
| | | |  
DB 5 CADGCRLL 12

RESULT 10  
US-08-867-941-28  
; Sequence 28, Application US/08867941  
; Patent No. 5977337  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena M  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Wang, Quijun  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Klein, Michel H  
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/867,941  
; FILING DATE: 03-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-867-941-28

Query Match 33.8%; Score 25; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 11e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 RYVOKGE 12  
| | | | | |  
DB 1 QYTRKGE 7

RESULT 11  
US-09-074-658-28  
; Sequence 28, Application US/09074658  
; Patent No. 6184371  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena M  
; APPLICANT: Run-Pan Du  
; APPLICANT: Quijun Wang  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Klein, Michel H  
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA

NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/074,658  
FILING DATE: 08-MAY-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-795  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-074-658-28

Query Match 33.8%; Score 25; DB 4; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.1e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 RYVKGE 12  
:|:|:|  
Db 1 QYTRKE 7

RESULT 12  
US-08-439-905-7  
Sequence 7, Application US/08439905  
Patent No. 5645815  
GENERAL INFORMATION:  
APPLICANT: Dean, Richard T.  
APPLICANT: Lister-James, John  
APPLICANT: Civitello, Edgar  
TITLE OF INVENTION: RADIOABLE COMPOUNDS FOR THROMBUS  
FILE REFERENCE: 104D2  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/439,905  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/044,825  
FILING DATE: 08-APR-1993  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/653,012  
FILING DATE: 08-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 07234/002001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: Each Cys residue is protected at the S by an acetamidomethyl group. The C-terminus is an amide (i.e., -CO-OH).  
US-08-439-905-7

Query Match 33.8%; Score 25; DB 1; Length 12;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RDGIR 6  
|||:|  
Db 2 RDGVR 6

RESULT 13  
US-09-100-536-8  
Sequence 8, Application US/09100536A  
Patent No. 6022520  
GENERAL INFORMATION:  
APPLICANT: Dean, Richard T.  
APPLICANT: Lister-James, John  
TITLE OF INVENTION: Radiolabeled Compounds for Thrombus Imaging  
FILE REFERENCE: 104D2  
CURRENT APPLICATION NUMBER: US/09/100,536A  
CURRENT FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 08/253,317  
EARLIER FILING DATE: 1994-06-03  
EARLIER APPLICATION NUMBER: 08/044,825  
EARLIER FILING DATE: 1993-04-08  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (10)  
OTHER INFORMATION: acetamidomethyl  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (12)  
OTHER INFORMATION: acetamidomethyl  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (12)  
OTHER INFORMATION: AMIDATION  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide  
US-09-100-536-8

Query Match 33.8%; Score 25; DB 3; Length 12;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RDGIR 6  
|||:|  
Db 2 RDGVR 6

RESULT 14  
US-09-100-537-8  
; Sequence 8, Application US/09100537A  
; Patent No. 6022857  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Richard T.  
; APPLICANT: Lister-James, John  
; TITLE OF INVENTION: Radiolabeled Compounds for Thrombus Imaging  
; FILE REFERENCE: 104D1  
; CURRENT APPLICATION NUMBER: US/09/100,537A  
; CURRENT FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 08/253,317  
; EARLIER FILING DATE: 1994-06-03  
; EARLIER APPLICATION NUMBER: 08/044,825  
; EARLIER FILING DATE: 1993-04-08  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (10)  
; OTHER INFORMATION: acetamidomethyl  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (12)  
; OTHER INFORMATION: acetamidomethyl  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (12)  
; OTHER INFORMATION: AMIDATION  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic  
; OTHER INFORMATION: peptide  
US-09-100-537-8

Query Match 33.8%; Score 25; DB 3; Length 12;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RDGIR 6  
|||:|  
Db 2 RDGVR 6

RESULT 15  
US-09-011-961-17  
; Sequence 17, Application US/09011961  
; Patent No. 6197536  
; GENERAL INFORMATION:  
; APPLICANT: STEINKUEHLER, Christian  
; APPLICANT: PESSI, Antonello  
; APPLICANT: BIANCHI, Elisabetta  
; APPLICANT: TALIANI, Marina  
; APPLICANT: TOMEL, Licia  
; APPLICANT: URBANI, Andrea  
; APPLICANT: DE FRANCESCO, Raffaele  
; APPLICANT: NARJES, Frank  
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND  
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE  
; TITLE OF INVENTION: HCV NS3 PROTEASE  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/011,961  
; FILING DATE: 23-FEB-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IT96/00163  
; FILING DATE: 20-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IT RM95A000573  
; FILING DATE: 22-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: YUN, Allen C.  
; REGISTRATION NUMBER: 37,971  
; REFERENCE/DOCKET NUMBER: STEINKUEHLER-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1  
; OTHER INFORMATION: /note= "Xaa is Ac-Glu"  
US-09-011-961-17

Query Match 33.8%; Score 25; DB 4; Length 12;  
Best Local Similarity 27.3%; Pred. No. 1.4e+02;  
Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRDGIRYVQKG 11  
| :|::|  
Db 2 CASHLPYIEQG 12

Search completed: February 21, 2002, 16:57:05  
Job time: 113 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:59:27 ; Search time 12.72 Seconds  
(without alignments)  
41.920 Million cell updates/sec

Title: US-08-753-851-10

Perfect score: 41

Sequence: 1 PSNPTDD 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 455

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	41.5	4	2 I40697	biotin A - Citrobacter
2	13	31.7	5	2 S51077	alpha-amylase - ri
3	13	31.7	5	2 PT0601	T-cell receptor be
4	13	31.7	6	2 PQ0008	angiotensin-conver
5	13	31.7	6	2 PT0533	T-cell receptor be
6	13	31.7	7	2 S16365	opacity protein P.
7	13	31.7	7	2 A15398	choline oxidase (E
8	13	31.7	7	2 B44787	caliIFMRamide 11
9	13	31.7	7	2 PT0628	T-cell receptor be
10	13	31.7	7	2 PT0576	T-cell receptor be
11	12	29.3	3	2 I78890	tyrosine protein k
12	12	29.3	4	2 S53508	starvation-induced
13	12	29.3	5	1 HOROHA	proctolin - Americ
14	12	29.3	5	2 PS0324	ribulose-bisphosph
15	12	29.3	5	2 A60411	proctolin - Atlant
16	12	29.3	5	2 PT0714	T-cell receptor be
17	12	29.3	5	2 PT0679	T-cell receptor be
18	12	29.3	6	2 PT0618	T-cell receptor be
19	12	29.3	6	2 PT0715	T-cell receptor be
20	12	29.3	7	2 S68004	hucolin, 75K chain
21	12	29.3	7	2 B39040	calsequestrin, fas
22	12	29.3	7	2 PT0722	T-cell receptor be
23	11	26.8	4	2 PT0675	T-cell receptor be
24	11	26.8	5	2 JT0520	Ig kappa chain V-I
25	11	26.8	6	4 S15596	orf 3 rara 5'-regi
26	11	26.8	7	1 A61324	dermorphin - Rohde
27	11	26.8	7	2 A61081	tryptophyllin, bas
28	11	26.8	7	2 PT0087	ribulose-bisphosph
29	11	26.8	7	2 PC1316	large granule L3 c

30 11 26.8 7 2 PT0515 T-cell receptor be  
31 11 26.8 7 2 A39690 neural cell adhesi  
32 10 24.4 4 2 A40135 branched-chain-am  
33 10 24.4 5 2 A37114 hypoxanthine phosph  
34 10 24.4 5 2 PT0695 T-cell receptor be  
35 10 24.4 7 2 A34026 acetylcholinestera  
36 10 24.4 7 2 A12016 formylglycinamide  
37 9 22.0 5 2 PT0656 T-cell receptor be  
38 9 22.0 5 2 PT0535 T-cell receptor be  
39 9 22.0 5 2 PT0561 T-cell receptor be  
40 9 22.0 6 2 A44916 mosquitoicidal toxi  
41 9 22.0 6 2 PT0280 Ig heavy chain CRD  
42 9 22.0 6 2 S29881 Na+/K+-exchanging  
43 9 22.0 6 2 PT0510 T-cell receptor be  
44 9 22.0 6 2 PT0512 T-cell receptor be  
45 9 22.0 6 2 PT0599 T-cell receptor be

#### ALIGNMENTS

RESULT 1  
I40697  
biotin A - Citrobacter freundii (fragment)  
C:Species: Citrobacter freundii  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 12-Aug-1996  
C:Accession: I40697  
R:Shiuan, D.; Campbell, A.  
Gene 67, 203-211, 1988  
A:Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter  
A:Reference number: I40697; MUID:89006280  
A:Accession: I40697  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-4 <RES>  
A:Cross-references: GB:M21922; NID:gl44434

Query Match 41.5%; Score 17; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TDD 7  
Db 2 TDD 4

RESULT 2  
S51077  
alpha-amylase - rice  
C:Species: Oryza sativa (rice)  
C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 01-Sep-1995  
C:Accession: S51077  
R:Terashima, M.; Kubo, A.; Suzawa, M.; Itoh, Y.; Katoh, S.  
Eur. J. Biochem. 226, 249-254, 1994  
A:Title: The roles of the N-linked carbohydrate chain of rice alpha-amylase in thermo  
A:Reference number: S51077; MUID:95045597  
A:Accession: S51077  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-5 <TER>

Query Match 31.7%; Score 13; DB 2; Length 5;  
Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SNPT 5  
Db 1 SNGT 4

RESULT 3

PT0601

T-cell receptor beta chain V-D-J region (120-2K) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0601; PT0617; PT0694  
 R:Feeney, A. J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601  
 A:Accession: PT0601  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-5 <FE2>  
 A:Experimental source: newborn thymus, strain BALB/c, clone 120-2K  
 A:Accession: PT0617  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-5 <FE3>  
 A:Experimental source: newborn thymus, strain BALB/c, 120-2CA  
 A:Accession: PT0694  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-5 <FE2>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-1H  
 C:Keywords: T-cell receptor

Query Match 31.7%; Score 13; DB 2; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7  
 :||  
 Db 3 SDD 5

RESULT 4

PQ0008  
 angiotensin-converting enzyme inhibitor (FLP-1) - common fig  
 N:Alternate names: ficus latex peptide 1  
 C:Species: Ficus carica (common fig)  
 C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 08-Dec-1995  
 C:Accession: PQ0008  
 R:Maruyama, S.; Miyoshi, S.; Tanaka, H.  
 Agric. Biol. Chem. 53, 2763-2767, 1989  
 A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.  
 A:Reference number: PQ0008  
 A:Accession: PQ0008  
 A:Molecule type: protein  
 A:Residues: 1-6 <MAR>  
 A:Experimental source: latex  
 C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 31.7%; Score 13; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4  
 :||  
 Db 3 NP 4

RESULT 5

PT0533  
 T-cell receptor beta chain V-D-J region (126-1AA) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0533  
 R:Feeney, A. J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601

A:Accession: PT0533  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-6 <FE2>  
 A:Experimental source: adult thymus, strain BALB/c  
 C:Keywords: T-cell receptor

Query Match 31.7%; Score 13; DB 2; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7  
 :||  
 Db 3 SDD 5

RESULT 6

SI6365  
 opacity protein P.IIf - Neisseria gonorrhoeae (strain FA1090) (fragment)  
 N:Alternate names: outer membrane protein P.IIf  
 C:Species: Neisseria gonorrhoeae  
 A:Variety: strain FA1090  
 C:Date: 19-Mar-1997 #sequence\_revision 17-Oct-1997 #text\_change 17-Oct-1997  
 C:Accession: SI6365  
 R:Barritt, D.S.; Schwalbe, R.S.; Klapper, D.G.; Cannon, J.G.  
 Infect. Immun. 55, 2026-2031, 1987  
 A:Title: Antigenic and structural differences among six proteins II expressed by a strain of Neisseria gonorrhoeae  
 A:Reference number: SI6360; MUID:87306843  
 A:Accession: SI6365  
 A:Molecule type: protein  
 A:Residues: 1-7 <BAR>

A:Experimental source: strain FA1090  
 A:Note: expression of opacity proteins is regulated by the number of translated repeats of repeats place the start codon in frame with the rest of the protein  
 C:Superfamily: opacity protein  
 C:Keywords: cell surface component; transmembrane protein  
 F.1-7/Product: opacity protein P.IIf (fragment) #status experimental <MAT>

Query Match 31.7%; Score 13; DB 2; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7  
 :||  
 Db 2 TED 4

RESULT 7

AI5398  
 choline oxidase (EC 1.1.3.17) - Alcaligenes sp. (tentative sequence) (fragment)  
 C:Species: Alcaligenes sp.  
 C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 31-Mar-2000  
 C:Accession: AI5398  
 R:Ohta-Fukuyama, M.; Miyake, Y.; Emi, S.; Yamano, T.  
 J. Biochem. 88, 197-203, 1980  
 A:Title: Identification and properties of the prosthetic group of choline oxidase from Alcaligenes sp.  
 A:Reference number: AI5398; MUID:81006769  
 A:Accession: AI5398  
 A:Molecule type: protein  
 A:Residues: 1-7 <OHT>  
 C:Keywords: oxidoreductase

Query Match 31.7%; Score 13; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4  
 :||  
 Db 2 NP 3

## RESULT 8

B44787  
Calliphormamide 11 - bluebottle fly (Calliphora vomitoria)  
C:Species: Calliphora vomitoria  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999  
C:Accession: B44787  
R:Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe  
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992  
A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi  
A:Reference number: A41978; MUID:92196111  
A:Accession: B44787  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <DUV>  
C:Keywords: amidated carboxyl end; neuropeptide  
F:7/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 31.7%; Score 13; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSN 3  
| |  
Db 1 PDN 3

## RESULT 9

PT0628  
T-cell receptor beta chain V-D-J region (111-118) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0628  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601  
A:Accession: PT0628  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-7 <FEE>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 31.7%; Score 13; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7  
| | |  
Db 3 SDD 5

## RESULT 10

PT0576  
T-cell receptor beta chain V-D-J region (141-148) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0576  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601  
A:Accession: PT0576  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-7 <FEE>  
A:Experimental source: day 19 fetal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 31.7%; Score 13; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7  
| | |  
Db 3 SDD 5

## RESULT 11

I78890  
tyrosine protein kinase - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I78890  
R:Chow, L.M.; Davidson, D.; Fournel, M.; Gosselin, P.; Lemieux, S.; Lyu, M.S.; Kozak,  
Oncogene 9, 3437-3448, 1994  
A:Title: Two distinct protein isoforms are encoded by nt, a csk-related tyrosine pro  
A:Reference number: I58407; MUID:95060800  
A:Accession: I78890  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3 <RES>  
A:Cross-references: GB:L33339; NID:g609536; PIDN:AAA64432.1; PID:g609538  
C:Genetics:  
A:Gene: p52ntk

Query Match 29.3%; Score 12; DB 2; Length 3;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PT 5  
| |  
Db 2 PT 3

## RESULT 12

S53508  
starvation-induced ribonuclease - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 07-May-1999  
C:Accession: S53508  
R:Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.  
Plant Mol. Biol. 27, 477-485, 1995  
A:Title: cDNA structure and regulatory properties of a family of starvation-induced r  
A:Reference number: S53506; MUID:95201242  
A:Accession: S53508  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-4 <KOE>

Query Match 29.3%; Score 12; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PT 5  
| |  
Db 2 PT 3

## RESULT 13

H0808A  
proctolin - American cockroach  
C:Species: Periplaneta americana (American cockroach)  
C:Date: 29-Jul-1981 #sequence\_revision 29-Jul-1981 #text\_change 23-Aug-1996  
C:Accession: A01644  
R:Starratt, A.N.; Brown, B.E.  
Life Sci. 17, 1253-1256, 1975  
A:Title: Structure of the pentapeptide proctolin, a proposed neurotransmitter in inse  
A:Reference number: A93048; MUID:76074708  
A:Accession: A01644

Search completed: February 21, 2002, 17:01:17  
Job time: 110 sec

A:Molecule type: protein  
A:Residues: 1-5 <STA>  
A:Note: the synthetic peptide had the same chromatographic, electrophoretic, and pharmacological properties as the natural peptide.  
R:O'Shea, M.; Adams, M.E.;  
Science 213, 567-569, 1981  
A:Title: Pentapeptide (proctolin) associated with an identified neuron.  
A:Reference number: A94260; MUID:81225865  
A:Contents: annotation; biological source  
C:Comment: This peptide is found in the lateral white neurons, which occur (in the cockroach) innervate the striated hindgut muscles in insects and stimulate contraction of these muscles.  
C:Superfamily: proctolin  
C:Keywords: neuropeptide

Query Match 29.3%; Score 12; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PT 5  
II  
DB 4 PT 5

RESULT 14  
PS0324  
ribulose-bisphosphate carboxylase activase III - rice (strain Nihonbare) (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 02-Jul-1998  
C:Accession: PS0324  
R:Tsuigita, A.  
submitted to JIPID, April 1993  
A:Reference number: PS0206  
A:Accession: PS0324  
A:Molecule type: protein  
A:Residues: 1-5 <TSU>  
A:Experimental source: leaf, chlorophyll

Query Match 29.3%; Score 12; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PT 5  
II  
DB 4 PT 5

RESULT 15  
A60411  
proctolin - Atlantic horseshoe crab  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 17-Mar-1999  
C:Accession: A60411  
R:Groomer, J.R.; Tillinghast, E.K.; Townley, M.A.; Vetrovs, A.; Watson III, W.H.; Hunt, D.  
Peptides 11, 205-211, 1990  
A:Title: Identification of proctolin in the central nervous system of the horseshoe crab.  
A:Reference number: A60411; MUID:90287800  
A:Accession: A60411  
A:Molecule type: protein  
A:Residues: 1-5 <GRO>  
C:Comment: This neuropeptide stimulates cardiac output and hindgut motility in the horseshoe crab.  
C:Keywords: neuropeptide

Query Match 29.3%; Score 12; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PT 5  
II  
DB 4 PT 5

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:01:02 ; Search time 10.07 Seconds  
(without alignments)  
25.487 Million cell updates/sec

Title: US-08-753-851-10

Perfect score: 41

Sequence: 1 PSNPTDD 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	46.3	6	1	TMOF_SARBU
2	17	41.5	5	1	BIOA_CITFR
3	17	41.5	5	1	BIOA_SALTY
4	14	34.1	6	1	VP19_HSV1K
5	13	31.7	7	1	ALL5_CARMA
6	13	31.7	7	1	CHOX_ALCSP
7	13	31.7	7	1	FARB_CALVO
8	12	29.3	5	1	PRTT_PERAM
9	12	29.3	7	1	UF04_MOUSE
10	8	19.5	3	1	THYL_PIG
11	8	19.5	4	1	DCML_PSECH
12	8	19.5	4	1	TUFT_HUMAN
13	8	19.5	5	1	BIOB_SALTY
14	8	19.5	5	1	TRM3_ECOLI
15	8	19.5	6	1	CIP1_MYED
16	8	19.5	7	1	ALL4_CARMA
17	8	19.5	7	1	FAR1_HELTI
18	8	19.5	7	1	FAR4_PANRE
19	8	19.5	7	1	MNPL_LEPDE
20	7	17.1	4	1	RM01_YEAST
21	7	17.1	5	1	BPF7_BOTIN
22	7	17.1	5	1	PAP2_PARMA
23	7	17.1	5	1	SUGA_ACHDO
24	7	17.1	5	1	TPIS_CANFA
25	7	17.1	6	1	ASP2_LACSN
26	7	17.1	6	1	CIP2_MYED
27	7	17.1	6	1	OVN_LEPDE
28	7	17.1	6	1	TRPL_PSEPU
29	7	17.1	6	1	UN06_CLOPA
30	7	17.1	7	1	ALL3_CARMA
31	7	17.1	7	1	CARP_MYED
32	7	17.1	7	1	GFRP_MOUSE
33	7	17.1	7	1	IGAO_DACDE

34	7	17.1	7	1	LANC_CARUI	P36960	carnobacter
35	7	17.1	7	1	MYOM_APLCA	P15513	aplysia cal
36	7	17.1	7	1	UN06_PINPS	P81675	pinus pinas
37	6	14.6	3	1	LUXE_VIBFI	P24272	vibrio fisc
38	6	14.6	4	1	ACH1_ACHFU	P35904	achatina fu
39	6	14.6	5	1	UXA4_CHLTR	P38005	chlamydia t
40	6	14.6	5	1	FARP_MONEX	P41966	moniezia ex
41	6	14.6	7	1	ALL7_CYDPO	P82158	cydia pomon
42	6	14.6	7	1	FAR1_ASCSU	P31889	ascaris suu
43	6	14.6	7	1	FAR1_PROCL	P38499	procambarus
44	6	14.6	7	1	FAR2_PROCL	P38498	procambarus
45	6	14.6	7	1	UF03_MOUSE	P38641	mus musculus

#### ALIGNMENTS

##### RESULT 1

TMOF\_SARBU STANDARD; PRT; 6 AA.  
AC P41495;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE TRYPsin-MODULATING COSTATIC FACTOR (TMOF).  
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Sarcophagidae; Sarcophaga.  
OX NCBI\_TaxID=7385;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Ovary;  
RX MEDLINE=94211930; PubMed=8159807;  
RA Bylemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,  
de Loof A.;  
RT "Sequencing and characterization of trypsin modulating oostatic  
factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria  
(Sarcophaga) bullata.";  
RL Regul. Pept. 50:61-72(1994).  
CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS  
CC IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN  
CC CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE  
CC DEVELOPMENT.  
CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR  
CC EPITHELIUM AFTER A BLOOD MEAL.  
KW Hormone.  
SQ SEQUENCE 6 AA; 695 MW; 61E72451B7642000 CRC64;

Query Match 46.3%; Score 19; DB 1; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NPTD 6  
!!!  
DB 1 NPTN 4

##### RESULT 2

BIOA\_CITFR STANDARD; PRT; 5 AA.  
AC P13071;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE  
DE (EC 2.6.1.62) (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA  
DE AMINOTRANSFERASE) (FRAGMENT).  
GN BIOA.  
OS Citrobacter freundli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Citrobacter.

```

OX NCBI_TaxID=546;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=89006280; PubMed=2971595;
RA Shivan D., Campbell A.;
RT "Transcriptional regulation and gene arrangement of Escherichia coli,
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL Gene 67:203-211(1988).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 8-AMINO-7-
CC OXONANATE = S-ADENOSYL-4-METHYLTHIO-2-OXOBUTANOATE +
CC 7,8-DIAMINONANATE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: BIOTIN BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M21922; -; NOT_ANNOTATED_CDS.
CC InterPro; IPR000954; Aminotran_3.
CC PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
CC Biotin biosynthesis; Transferase; Aminotransferase;
CC Pyridoxal phosphate. 5
CC NON_TER 5
CC SEQUENCE 5 AA; 582 MW; 6AAAB1BIA6F00000 CRC64;

Query Match 41.5%; Score 17; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. le+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7
DB 3 TDD 5

RESULT 3
BIOA_SALTY STANDARD; PRT; 5 AA.
ID BIOA_SALTY
AC P12677;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ADENOSYLMETHIONINE-8-AMINO-7-OXONANATE AMINOTRANSFERASE
DE (EC 2.6.1.62) (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA
DE AMINOTRANSFERASE) (FRAGMENT).
GN BIOA.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=89006280; PubMed=2971595;
RA Shivan D., Campbell A.;
RT "Transcriptional regulation and gene arrangement of Escherichia coli,
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL Gene 67:203-211(1988).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 8-AMINO-7-
CC OXONANATE = S-ADENOSYL-4-METHYLTHIO-2-OXOBUTANOATE +
CC 7,8-DIAMINONANATE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: BIOTIN BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M21923; -; NOT_ANNOTATED_CDS.
CC StyGene; SG10026; bioA.
CC InterPro; IPR000954; Aminotran_3.
CC PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
CC Biotin biosynthesis; Transferase; Aminotransferase;
CC Pyridoxal phosphate. 5
CC NON_TER 5
CC SEQUENCE 5 AA; 582 MW; 6AAAB1BIA6F00000 CRC64;

Query Match 41.5%; Score 17; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. le+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TDD 7
DB 3 TDD 5

RESULT 4
VP19_HSV1K STANDARD; PRT; 6 AA.
ID VP19_HSV1K
AC P23210;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE CAPSID ASSEMBLY AND DNA MATURATION PROTEIN (VIRION PROTEIN UL38)
DE (CAPSID PROTEIN VP19C) (FRAGMENT).
GN UL38.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10306;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=91101287; PubMed=1846198;
RA Flanagan W.M., Papavassiliou A.G., Rice M., Hecht L.B.,
RA Silverstein S., Wagner E.K.;
RT "Analysis of the herpes simplex virus type 1 promoter controlling the
RT expression of UL38, a true late gene involved in capsid assembly.";
RL J. Virol. 65:769-786(1991).
CC -1- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC EMBEDDED. BINDS DNA.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP19C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57646; AAA45830.1; -;
CC Capsid assembly; Coat protein; DNA-binding.
CC NON_TER 6
CC SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 34.1%; Score 14; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. le+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SNP 4

```



Db 3 TNP 5

RESULT 5  
ALL5\_CARMA  
ID ALL5\_CARMA STANDARD; PRT; 7 AA.  
AC P81808;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CARCINUSTATIN 5.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunioidea; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestros J.-L., Scott A.G., Jaros P.P.,  
Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the  
allatostatin superfamily in the shore crab Carcinus maenas.";  
RL Eur. J. Biochem. 250:727-734(1997).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation; Multigene family.  
FT MOD\_RES 7  
SQ SEQUENCE 7 AA; 781 MW; 672879CDBC476420 CRC64;

Query Match 31.7%; Score 13; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4  
||  
Db 1 NP 2

RESULT 6  
CHOX\_ALCSP  
ID CHOX\_ALCSP STANDARD; PRT; 7 AA.  
AC P16101;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DE CHOLINE OXIDASE (EC 1.1.3.17) (FRAGMENT).  
OS Alcaligenes sp.  
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
OC Alcaligenes.  
OX NCBI\_TaxID=512;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=81006769; PubMed=6997283;  
RA Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;  
RT "Identification and properties of the prosthetic group of choline  
oxidase from Alcaligenes sp.";  
RL J. Biochem. 88:197-203(1980).  
CC -!- CATALYTIC ACTIVITY: CHOLINE + O(2) = BETAINE ALDEHYDE + H(2)O(2).  
DR PIR: A15398; A15398.  
KW Oxidoreductase.  
FT NON\_TER 7  
SQ SEQUENCE 7 AA; 839 MW; 7415B1E457644ACO CRC64;

Query Match 31.7%; Score 13; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4

Db 2 NP 3

RESULT 7  
FARB\_CALVO  
ID FARB\_CALVO STANDARD; PRT; 7 AA.  
AC P41866;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CALLIFMRFAMIDE 11.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Thoracic ganglion;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
neuropeptides (designated callifMRFamides) from the blowfly  
Calliphora vomitoria.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
DR PIR: B44787; B44787.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7  
SQ SEQUENCE 7 AA; 926 MW; 69D40699C44AB700 CRC64;

Query Match 31.7%; Score 13; DB 1; Length 7;  
Best Local Similarity 66.7%; Pred. No. 1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSN 3  
||  
Db 1 PDN 3

RESULT 8  
PRCT\_PERAM  
ID PRCT\_PERAM STANDARD; PRT; 5 AA.  
AC P01373;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE PROCTOLIN.  
OS Periplaneta americana (American cockroach),  
OS Limulus polyphemus (Atlantic horseshoe crab), and  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blattodea; Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978, 6850, 6759;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=P. americana;  
RX MEDLINE=76074708; PubMed=576;  
RA Starratt A.N., Brown B.E.;  
RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter  
in insects.";  
RL Life Sci. 17:1253-1256(1975).  
RN [2]  
RP BIOLOGICAL SOURCE.  
RC SPECIES=P. americana;  
RX MEDLINE=81225865; PubMed=6113690;  
RA O'Shea M., Adams M.E.;  
RT "Pentapeptide (proctolin) associated with an identified neuron.";

```

RL Science 213:567-569(1981).
RN [3]
RP SEQUENCE.
RC SPECIES=L.polyphemus;
RX MEDLINE=90287800; PubMed=2356151;
RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA Shabanowitz J.;
RT "Identification of proctolin in the central nervous system of the
RT horseshoe crab, Limulus polyphemus.";
RL Peptides 11:205-211(1990).
RN [4]
RP SEQUENCE.
RC SPECIES=C.maenas; PubMed=2872661;
RX MEDLINE=86232789; PubMed=2872661;
RA Stangier J., Dirksen H., Keller R.;
RT "Identification and immunocytochemical localization of proctolin in
RT pericardial organs of the shore crab, Carcinus maenas.";
RL Peptides 7:67-72(1986).
CC -I- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY.
CC MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC -I- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC THE CRAB PERICARDIAL ORGANS.
DR PIR; A01644; HOROHA.
DR PIR; A60411; A60411.
KW Neuropeptide.
SQ SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;

Query Match 29.3%; Score 12; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 PT 5
Db 4 PT 5

RESULT 9
UF04_MOUSE
ID UF04_MOUSE STANDARD; PRT; 7 AA.
AC P38642;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF FIBROBLASTS (P46) (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast; PubMed=7523108;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.0, ITS MW IS: 46 KDA.
FT NON_TER
RN 7
SQ SEQUENCE 7 AA; 766 MW; 68640AB777632700 CRC64;

Query Match 29.3%; Score 12; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSNP 4
Db 1 PKPP 4

```

---

```

RESULT 10
THYL_PIG
ID THYL_PIG STANDARD; PRT; 3 AA.
AC P01151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE THYROLIBERIN (THYROTROPIN RELEASING HORMONE) (TRH).
OS Sus scrofa (Pig); Ovis aries (Sheep).
OS Bombina orientalis (Oriental fire-bellied toad), and
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823, 9940, 8346, 8316;
RN [1]
RP SEQUENCE.
RC SPECIES=Pig; TISSUE=Hypothalamus;
RX MEDLINE=70136150; PubMed=4984938;
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
RT "Structure of porcine thyrotropin releasing hormone.";
RL Biochemistry 9:1103-1106(1970).
RN [2]
RP SEQUENCE.
RC SPECIES=Sheep; TISSUE=Hypothalamus;
RX MEDLINE=70039904; PubMed=4982117;
RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
RT "The identity of chemical and hormonal properties of the thyrotropin
RT releasing hormone and pyroglutanyl-histidyl-proline amide.";
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
RN [3]
RP SEQUENCE.
RC SPECIES=Sheep; TISSUE=Hypothalamus;
RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
RA Ward D.N.;
RT "The elucidation of the primary structure of the hypothalamic thyroid
RT stimulating hormone releasing factor of ovine origin by means of mass
RT spectrometry.";
RL Org. Mass Spectrom. 5:221-228(1971).
RN [4]
RP SYNTHESIS.
RC SPECIES=Sheep; TISSUE=Hypothalamus;
RX MEDLINE=70163386; PubMed=4985794;
RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RA Guillemin R.;
RT "Characterization of ovine hypothalamic hypophysiotropic
RT TSH-releasing factor.";
RL Nature 226:321-325(1970).
RN [5]
RP SEQUENCE.
RC SPECIES=B.orientalis; TISSUE=Skin;
RX MEDLINE=76138399; PubMed=815011;
RA Yasuhara T., Nakajima T.;
RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RL Chem. Pharm. Bull. 23:3301-3303(1975).
RN [6]
RP SEQUENCE.
RC SPECIES=N.viridescens;
RX MEDLINE=75035605; PubMed=4214528;
RA Grimm-Joetgens Y., McKelvy J.F.;
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT viridescens) brain in vitro. Isolation and characterization of
RT thyrotropin releasing factor.";
RL J. Neurochem. 23:471-478(1974).
CC -I- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
DR PIR; A01415; RHPGT.
DR PIR; A93750; RHST.
DR PIR; A90919; RHTDIO.
DR PIR; A92971; A92971.
KW Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
MOD_RES 3 3 AMIDATION.

```

SQ SEQUENCE 3 AA; 380 MW; 7761F6B0000000000000 CRC64;

Query Match 19.5%; Score 8; DB 1; Length 3;

Best Local Similarity 50.0%; Pred. No. 1e+05;

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4

:|

Db 2 HP 3

RESULT 11

DCML\_PSECH

ID DCML\_PSECH STANDARD; PRT; 4 AA.

AC P19916;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-DEC-1992 (Rel. 24, Last annotation update)

DE CARBON MONOXIDE DEHYDROGENASE LARGE CHAIN (EC 1.2.99.2) (FRAGMENT).

OS Pseudomonas carboxydohydrogena.

OC Bacteria; Proteobacteria.

OX NCBI\_TaxID=290;

RN [1]

RP SEQUENCE.

RX MEDLINE=90055678; PubMed=2818128;

RA Kraut M., Hugendieck I., Herwig S., Meyer O.;

RT "Homology and distribution of CO dehydrogenase structural genes in

RT carboxydohydrophobic bacteria.";

RL Arch. Microbiol. 152:335-341(1989).

CC -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED

CC ACCEPTOR.

CC -!- COFACTOR: MOLYBDENUM.

CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND

CC SMALL.

DR PIR; P10140; P10140.

KW Oxidoreductase; Molybdenum.

FT NON\_TER 4 4

SQ SEQUENCE 4 AA; 441 MW; 7761E876F00000000000 CRC64;

Query Match

Best Local Similarity 19.5%; Score 8; DB 1; Length 4;

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4

:|

Db 3 HP 4

RESULT 12

TUFT\_HUMAN

ID TUFT\_HUMAN STANDARD; PRT; 4 AA.

AC P01858;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PHAGOCYTOSIS-STIMULATING PEPTIDE (TUFTSIN).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=72187087; PubMed=4112769;

RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;

RT "The characteristics, isolation and synthesis of the phagocytosis

RT stimulating peptide tuftsins.";

RL Biochem. Biophys. Res. Commun. 47:172-179(1972).

RN [2]

RP IMMUNOGLOBULIN CLASS.

RX MEDLINE=68091045; PubMed=4169272;

RA Fidalgo B.V., Najjar V.A.;

RT "The physiological role of the lymphoid system. VI. The stimulatory effect of leucophilic gamma globulin (leucokinin) on the phagocytic activity of human polymorphonuclear leucocyte.";

RL Biochemistry 6:3386-3392(1967).

CC -!- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN. TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC ACTIVITY OF NEUTROPHILS.

DR PIR; A02147; A02147.

DR MIM; 191150; -.

SQ SEQUENCE 4 AA; 501 MW; 74176321C000000000 CRC64;

Query Match 19.5%; Score 8; DB 1; Length 4;

Best Local Similarity 33.3%; Pred. No. 1e+05;

Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SNP 4

:|

Db 1 TKP 3

RESULT 13

BIOB\_SALTY

ID BIOB\_SALTY STANDARD; PRT; 5 AA.

AC P12678;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE BIOTIN SYNTHASE (EC 2.8.1.6) (BIOTIN SYNTHETASE) (FRAGMENT).

GN BIOB.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

OX NCBI\_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89006280; PubMed=2971595;

RA Shiu D., Campbell A.;

RT "Transcriptional regulation and gene arrangement of Escherichia coli,

RT Citrobacter freundii and Salmonella typhimurium biotin operons.";

RL Gene 67:203-211(1988).

CC -!- CATALYTIC ACTIVITY: DETHIOBIOTIN + (S) - BIOTIN.

CC -!- PATHWAY: LAST STEP IN BIOTIN BIOSYNTHESIS PATHWAY.

CC -!- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES

CC FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M21923; -; NOT\_ANNOTATED\_CDS.

DR StyGene; SG10027; bioB.

KW Biotin biosynthesis; Iron-sulfur; Transferase.

FT NON\_TER 5 5

SQ SEQUENCE 5 AA; 611 MW; 7761F40DD6F000000000 CRC64;

Query Match

Best Local Similarity 19.5%; Score 8; DB 1; Length 5;

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4

:|

Db 4 HP 5

RESULT 14

```

TRM3_ECOLI
ID TRM3_ECOLI STANDARD; PRT; 5 AA.
AC P13973;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TRAM PROTEIN (FRAGMENT).
GN TRAM.
OS Escherichia coli.
OG Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88227859; PubMed=2836369;
RA Inamoto S., Yoshioka Y., Ohtsubo E.;
RT "Identification and characterization of the products from the traJ
RT and traY genes of plasmid R100.";
RL J. Bacteriol. 170:2749-2757(1988).
CC -1- FUNCTION: TRANSFER GENE PROTEIN. IS INVOLVED IN THE CONJUGATION
CC PROCESS OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: TO TRAM PROTEIN OF OTHER PLASMIDS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20941; -: NOT_ANNOTATED_CDS.
DR PIR: A32014; A32014.
KW Conjugation; Plasmid; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 5 AA; 634 MW; 6B1B1AAA443500000 CRC64;

Query Match 19.5%; Score 8; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. le+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 DD 7
DB 3 DE 4

RESULT 15
CIPI_MYTED
ID CIPI_MYTED STANDARD; PRT; 6 AA.
AC P13736;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CONTRACTION-INHIBITING PEPTIDE I (MIP I).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE=Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -1- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
CC MUSCLES.
CC -1- SIMILARITY: TO MIP II.
DR PIR: A27696; A27696.
KW Hormone; Amidation.

```

FT MOD\_RES 6 6 AMIDATION.  
SQ SEQUENCE 6 AA; 637 MW; 72C9C6B775B81000 CRC64;

Query Match 19.5%; Score 8; DB 1; Length 6;  
Best Local Similarity 50.0%; Pred. No. le+05;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NP 4  
DB 2 SP 3

Search completed: February 21, 2002, 17:04:00  
Job time: 178 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:00:42 ; Search time 22.11 Seconds  
(without alignments)  
46.310 Million cell updates/sec

Title: US-08-753-851-10  
Perfect score: 41  
Sequence: 1 PSNPTDD 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 65

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_17.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_invertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	36.6	7	12 Q07624	Q07624 rous sarcom
2	12	29.3	7	10 Q49223	Q49223 glycine max
3	11	26.8	7	10 P93233	P93233 lycopersico
4	11	26.8	7	11 Q63668	Q63668 rattus norv
5	10	24.4	7	12 Q65578	Q65578 bovine herp
6	9	22.0	7	2 P70804	P70804 azotobacter
7	8	19.5	5	13 P82099	P82099 litoria rub
8	8	19.5	7	6 Q28742	Q28742 oryctolagus
9	7	17.1	5	10 Q99007	Q99007 hordeum vul
10	7	17.1	6	13 P82096	P82096 litoria rub
11	7	17.1	7	2 Q54248	Q54248 streptomyce
12	7	17.1	7	2 O50556	O50556 actinobacil
13	7	17.1	7	2 Q47505	Q47505 escherichia
14	7	17.1	7	8 P92421	P92421 psathyrosta
15	7	17.1	7	8 P92385	P92385 hordeum mar
16	7	17.1	7	8 P92210	P92210 agropyron c
17	7	17.1	7	8 P92214	P92214 amblyopyrum
18	7	17.1	7	8 P92218	P92218 australopyr
19	7	17.1	7	8 P92221	P92221 bromus iner

20	7	17.1	7	8 P92226	P92226 crithopsis
21	7	17.1	7	8 P92372	P92372 haynaldia v
22	7	17.1	7	8 P92381	P92381 hordeum bra
23	7	17.1	7	8 P92387	P92387 henrardia p
24	7	17.1	7	8 P92390	P92390 heteranthe
25	7	17.1	7	8 P92393	P92393 hordeum vul
26	7	17.1	7	8 P92425	P92425 pseudoroegn
27	7	17.1	7	8 P92427	P92427 peridictyon
28	7	17.1	7	8 P92430	P92430 aegilops ta
29	7	17.1	7	8 P92442	P92442 taeniatheru
30	7	17.1	7	8 P92440	P92440 thinopyrum
31	7	17.1	7	8 P92403	P92403 lophopyrum
32	7	17.1	7	8 O98866	O98866 spinacia ol
33	7	17.1	7	12 Q66113	Q66113 cherry leaf
34	7	17.1	7	13 Q42564	Q42564 fugu rubrip
35	7	17.1	7	13 P82065	P82065 litoria rub
36	7	17.1	7	13 P82101	P82101 litoria rub
37	6	14.6	4	11 Q08433	Q08433 rattus norv
38	6	14.6	5	13 P82070	P82070 litoria rub
39	6	14.6	7	2 O07354	O07354 synechococc
40	6	14.6	7	2 O34028	O34028 pseudomonas
41	6	14.6	7	11 Q63480	Q63480 rattus norv
42	6	14.6	7	11 O55184	O55184 rattus norv
43	6	14.6	7	12 Q67113	Q67113 influenza a
44	5	12.2	5	13 P82072	P82072 litoria rub
45	5	12.2	5	13 P82073	P82073 litoria rub

#### ALIGNMENTS

RESULT 1

Q07624 ID Q07624 PRELIMINARY; PRT; 7 AA.  
AC Q07624;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE HYPOTHETICAL PROTEIN (FRAGMENT).  
OS Rous sarcoma virus  
OC Viruses; Retrovird viruses; Retroviridae; Alpharetrovirus.  
OC NCBI\_TaxID=11886;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PRAGUE C;  
RX MEDLINE=93010967; PubMed=1327749;  
RA Donze O., Spahr P.F.;  
RT "Role of the open reading frames of Rous sarcoma virus leader RNA in  
RT translation and genome packaging.";  
RL EMBO J. 11:3747-3757(1992).  
DR EMBL; X67587; CAA47862.1; -  
KW Hypothetical protein.  
FT NON\_TER 7  
SQ SEQUENCE 7 AA; 672 MW; 776045A7687DB6F0 CRC64;

Query Match 36.6%; Score 15; DB 12; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.7e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PSNP 4

Db 4 PSIP 7

RESULT 2

O49223 ID O49223 PRELIMINARY; PRT; 7 AA.  
AC O49223;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE HMG-1-LIKE PROTEIN (FRAGMENT).

OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroids I; Fabales; Fabaceae; Papilionoideae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. ESSEX; TISSUE=ROOTS;  
 RX MEDLINE=91367679; PubMed=1891369;  
 RA Laux T., Goldberg R.B.;  
 RT "A plant DNA binding protein shares highly conserved sequence motifs  
 with HMG-box proteins.";  
 RL Nucleic Acids Res. 19:4769-4769(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. ESSEX; TISSUE=ROOTS;  
 RA Mahalingam R., Knap H.T.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF047050; AAC03556.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

Query Match 29.3%; Score 12; DB 10; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DD 7  
 Db 5 DD 6

RESULT 3  
 P93233  
 ID P93233 PRELIMINARY; PRT; 7 AA.  
 AC P93233;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE (EC 4.4.1.14) (ACC  
 DE SYNTHASE) (FRAGMENT).  
 GN LE-ACS1B.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97351561; PubMed=9207843;  
 RA Ostler J.H., Olson D.C., Shiu O.Y., Yang S.F.;  
 RT "Differential induction of seven 1-aminocyclopropane-1-carboxylate  
 synthase genes by elicitor in suspension cultures of tomato  
 (Lycopersicon esculentum).";  
 RL Plant Mol. Biol. 34:275-286(1997).  
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE = 1-  
 CC AMINOCYCLOPROPANE-1-CARBOXYLATE + METHYLTHIOADENOSINE.  
 CC -1- COFACTOR: PYRIDOXAL-PHOSPHATE.  
 DR EMBL; U75692; AAC49682.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match 26.8%; Score 11; DB 10; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 4.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SNP 4  
 Db 1  
 3. SRP 3

RESULT 4  
 Q63668  
 ID Q63668 PRELIMINARY; PRT; 7 AA.  
 AC Q63668;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE VASOPRESSIN V2 RECEPTOR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;  
 RX MEDLINE=95396550; PubMed=7667072;  
 RA Mandon B., Bellanger A.C., Elalouf J.M.;  
 RT "Inverse PCR-mediated cloning of the promoter for the rat vasopressin  
 V2 receptor gene.";  
 RL Pflügers Arch. 430:12-18(1995).  
 DR EMBL; X83264; CAA58237.1; -;  
 SQ SEQUENCE 7 AA; 703 MW; 75A767287DC6D6F0 CRC64;

Query Match 26.8%; Score 11; DB 11; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PS 2  
 Db 6 PS 7

RESULT 5  
 Q65578  
 ID Q65578 PRELIMINARY; PRT; 7 AA.  
 AC Q65578;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL PROTEIN (FRAGMENT).  
 OS Bovine herpesvirus 1.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=10320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COOPER;  
 RX MEDLINE=95313343; PubMed=7793062;  
 RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,  
 RA Letchworth G.J., Schwyzler M.;  
 RT "Nucleotide sequence analysis of a 30-kb region of the bovine  
 herpesvirus 1 genome which exhibits a colinear gene arrangement with  
 the UL21 to UL4 genes of herpes simplex virus.";  
 RL Virology 210:100-108(1995).  
 DR EMBL; Z48053; CAA88130.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 758 MW; 6DD33455B1F1B1C0.CRC64;

Query Match 24.4%; Score 10; DB 12; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SN 3  
 Db 4 SN 5

RESULT 6  
 P70804  
 ID P70804 PRELIMINARY; PRT; 7 AA.

P70804;  
 AC 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE ALGG GENE (FRAGMENT).  
 GN ALGT.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Azotobacter.  
 OX NCBI\_TaxID=354;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=E;  
 RX MEDLINE=96427318; PubMed=8830682;  
 RA Rehm B.H.A., Ertesvag H., Valla S.;  
 RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is  
 part of an alg gene cluster physically organized in a manner similar  
 to that in Pseudomonas aeruginosa.";  
 RL J. Bacteriol. 178:5884-5889(1996).  
 DR EMBL: X87973; CAA61230.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;  
 Query Match 22.0%; Score 9; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 SNPT 5  
 Db 4 SSST 7  
 RESULT 7  
 ID P82099 PRELIMINARY; PRT; 5 AA.  
 AC P82099;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE ELECTRIN 3.  
 OS Litoria rubella (Desert tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
 OC Litoria.  
 OX NCBI\_TaxID=104895;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=SKIN SECRETION;  
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
 RT "Peptides from the skin glands of the Australian buzzing tree frog  
 Litoria electrica. Comparison with the skin peptides from Litoria  
 rubella.";  
 RL Aust. J. Chem. 52:0-0(1999).  
 KW Amphibian skin; Amidation.  
 FT MOD\_RES 5  
 SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;  
 Query Match 19.5%; Score 8; DB 13; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 NP 4  
 Db 3 HP 4  
 RESULT 8  
 ID Q28742 PRELIMINARY; PRT; 7 AA.  
 AC Q28742;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE ALPHA-MYOSIN HEAVY CHAIN (FRAGMENT).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84221901; PubMed=6328491;  
 RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.J., Jokovic S.,  
 RA Rabinowitz M.;  
 RT "Characterization of genomic clones specifying rabbit alpha- and beta-  
 ventricular myosin heavy chains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).  
 DR EMBL: K01698; AAA31415.1; -;  
 KW Myosin.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 916 MW; 6B1B1AA1E69326B0 CRC64;  
 Query Match 19.5%; Score 8; DB 6; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 DD 7  
 Db 5 DE 6

RESULT 9  
 Q99007 PRELIMINARY; PRT; 5 AA.  
 ID Q99007;  
 AC Q99007;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE ALPHA-AMYLASE (EC 3.2.1.1) (FRAGMENT).  
 GN AMY1.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HIMALAYA; TISSUE=ALEURONE LAYER;  
 RX MEDLINE=91329704; PubMed=1831055;  
 RA Jacobsen J.V., Close T.J.;  
 RT "Control of transient expression of chimaeric genes by gibberellic  
 acid and abscisic acid in protoplasts prepared from mature barley  
 aleurone layers.";  
 RL Plant Mol. Biol. 16:713-721(1991).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.  
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA-AMYLASE IN  
 BARLEY.

EMBL: X54643; CAA38455.1; -;  
 DR Hydrolyase; Glycosidase; Carbohydrate metabolism; Seed; Germination;  
 KW Calcium; Multigene family.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;  
 Query Match 17.1%; Score 7; DB 10; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 SN 3  
 Db 2 AN 3



```

RESULT 10
P82096 ID P82096 PRELIMINARY; PRT; 6 AA.
AC P82096;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE ELECTRIN 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN SECRETION;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella."
RL Aust. J. Chem. 52:0-0(1999).
KW Amphibian skin; Amidation.
FT MOD_RES 6
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 17.1%; Score 7; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 P 1
Db 3 P 3

RESULT 11
Q54248 ID Q54248 PRELIMINARY; PRT; 7 AA.
AC Q54248;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)
DE SECY & ADK GENES (FRAGMENT).
GN RPL0.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RA Foehtling S., Piepersberg W., Wehmeier U.F.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X95915; CAA65160.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 17.1%; Score 7; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 4.7e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TD 6
Db 4 TE 5

RESULT 12
O50556 ID O50556 PRELIMINARY; PRT; 7 AA.
AC O50556;
DT 01-JUN-1998 (TREMBlrel. 06, Created)

```

```

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GLYA (FRAGMENT).
GN GLYA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33384; PubMed=8751884;
RX Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RA "cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans."
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 17.1%; Score 7; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 P 1
Db 4 P 4

RESULT 13
Q47505 ID Q47505 PRELIMINARY; PRT; 7 AA.
AC Q47505;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE PLASMID PMCC7 MCCA,B,C,D,E,F GENES.
GN MCCA.
OS Escherichia coli.
OC Plasmid pmCC7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96099297; PubMed=8522520;
RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;
RT "Structure and organization of plasmid genes required to produce the
RT translation inhibitor microcin C7."
RL J. Bacteriol. 177:7131-7140(1995).
DR EMBL; X57583; CAA40808.1; -.
KW Plasmid.
SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 17.1%; Score 7; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 4.7e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SN 3
Db 6 AN 7

RESULT 14
P92421 ID P92421 PRELIMINARY; PRT; 7 AA.
AC P92421; P92419;
DT 01-MAY-1997 (TREMBlrel. 03, Created)

```

DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last annotation update)  
 DE RIBOSOMAL PROTEIN 11 (FRAGMENT).  
 GN RPS11.  
 OS Psathyrostachys fragilis.  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Psathyrostachys.  
 OX NCBI\_TaxID=37729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H4372, AND H917; TISSUE=LEAVES;  
 RA Petersen G., Seberg O.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; 277753; CAB01337.1; -.  
 DR EMBL; 277752; CAB01334.1; -.  
 KW Ribosomal protein; Chloroplast.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 17.1%; Score 7; DB 8; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 P 1  
 Db 1 P 1

RESULT 15  
 P92385 PRELIMINARY; PRT; 7 AA.  
 AC P92385; P92383;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last annotation update)  
 DE RIBOSOMAL PROTEIN 11 (FRAGMENT).  
 GN RPS11.  
 OS Hordeum marinum (Seaside barley).  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H299, AND H801; TISSUE=LEAVES;  
 RA Petersen G., Seberg O.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; 277763; CAB01367.1; -.  
 DR EMBL; 277762; CAB01364.1; -.  
 KW Ribosomal protein; Chloroplast.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 894 MW; 6734140333277700 CRC64;

Query Match 17.1%; Score 7; DB 8; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 P 1  
 Db 1 P 1

Search completed: February 21, 2002, 17:03:44  
 Job time: 182 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:58:47 ; Search time 23.47 Seconds  
(without alignments)  
22.093 Million cell updates/sec

Title: US-08-753-851-10

Perfect score: 41

Sequence: 1 PSNPTDD 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 47201

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	63.4	7	14 AAR34137	Antigenic peptide
2	24	58.5	7	20 AAY17785	B. thuringiensis c
3	24	58.5	7	20 AAY01210	Conserved peptide
4	24	58.5	7	21 AAY82403	B. thuringiensis c
5	24	58.5	7	22 AAU05017	Peptide encoded by
6	24	58.5	7	22 AAU05024	peptide encoded by
7	23	56.1	7	22 AAE06219	Ovine Calpha-s pep
8	22	53.7	7	19 AAW78635	SH2 domain binding
9	22	53.7	7	20 AAW95063	urokinase peptide
10	21	51.2	6	19 AAW47462	Prothrombin/thromb
11	21	51.2	7	17 AAR99608	RNA polymerase II

12	21	51.2	7	18 AAW18362	RNA polymerase II
13	21	51.2	7	20 AAY40791	Amino acid loop 7
14	21	51.2	7	20 AAY17818	B. thuringiensis c
15	21	51.2	7	20 AAY17819	Peptide SEQ ID NO:
16	21	51.2	7	20 AAY17786	B. thuringiensis c
17	21	51.2	7	21 AAB30131	Scaffold protein S
18	21	51.2	7	21 AAY82402	B. thuringiensis c
19	21	51.2	7	21 AAY82404	B. thuringiensis c
20	21	51.2	7	21 AAY82429	B. thuringiensis c
21	20	48.8	5	12 AAR12725	Pentapeptide paral
22	20	48.8	5	15 AAR51589	Pentapeptide #
23	20	48.8	5	16 AAR69957	Mimotopic mimotop
24	20	48.8	5	17 AAR98703	Pentameric peptide
25	20	48.8	6	12 AAR13940	Peptide 83 from 88
26	20	48.8	7	15 AAR50226	Semi-synthetic N-a
27	20	48.8	7	17 AAR99614	Sequence of linker
28	20	48.8	7	17 AAR99615	RNA polymerase II
29	20	48.8	7	17 AAR99618	RNA polymerase II
30	20	48.8	7	18 AAW18373	RNA polymerase II
31	20	48.8	7	18 AAW18369	RNA polymerase II
32	20	48.8	7	18 AAW18370	RNA polymerase II
33	20	48.8	7	20 AAW95064	Urokinase peptide
34	20	48.8	7	22 AAB92404	Miscellaneous pept
35	19	46.3	4	18 AAY08045	Biotin derivative
36	19	46.3	4	20 AAW83471	Human growth hormo
37	19	46.3	5	18 AAY08046	Biotin derivative
38	19	46.3	6	14 AAR39587	Diuretic hormone b
39	19	46.3	6	15 AAR57344	Peptide fragment o
40	19	46.3	6	16 AAR79351	Bovine DRADA amino
41	19	46.3	6	16 AAR82555	Neb-TMOF, peptide
42	19	46.3	6	17 AAR90507	Hybridoma ATCC HB-
43	19	46.3	6	17 AAR96233	Peptide hormone, N
44	19	46.3	6	18 AAY08047	Biotin derivative
45	19	46.3	6	18 AAW31260	Bovine 88 kD DRADA

#### ALIGNMENTS

RESULT 1  
AAR34137  
ID AAR34137 standard; peptide; 7 AA.  
XX AAR34137;  
XX 19-JUL-1993 (first entry)  
DT Antigenic peptide from liver cell regulation protein.  
DE Monoclonal antibody DSM ACC 2011; hepatocyte; sinusoidal pole;  
KW cell regulatory factor; MAb L8; cell adhesion molecule;  
KW cell-cell interaction; glycoprotein.  
XX Rattus.  
OS  
FH Key Location/Qualifiers  
FT Misc-difference 4 /note= "printed in parentheses,  
FT significance not defined"

XX WO9306134-A.  
XX 01-APR-1993.  
XX 16-SEP-1992; 92WO-FR00867.  
XX 16-SEP-1991; 91FR-0011389.  
XX (INRM ) INSERM INST NAT SANTE & RECH MED.  
XX Corlu A, Guillozo C, Kneip B;  
XX WPI; 1993-117474/14.

XX Protein(s) useful as cell regulation factor(s) - specifically  
 PT reactive with antibodies from animal immunised with rat liver  
 PT epithelial cells  
 XX  
 PS Claim 11; Page 40; 72pp; French.  
 XX  
 CC This peptide fragment is derived from a cell regulation protein  
 CC having an apparent molecular weight of 85,000 daltons, with a  
 CC pI of 4.9 to 5.1. The peptide is antigenic and specifically reacts  
 CC with a monoclonal antibody obtained by immunising an animal with  
 CC rat liver epithelial cells, followed by fusing and cloning. The  
 CC liver cell regulatory factor is localised to the membrane.  
 CC Particularly at the sinusoidal pole, of normal adult hepatocytes.  
 CC The proteins of the invention, as well as being expressed by  
 CC hepatocytes, are expressed by epithelial cells of the biliary ducts,  
 CC endothelial cells, Ito cells and macrophages, by Sertoli cells, by  
 CC spermatocytes at a specific developmental stage (end of the leptotene  
 CC and zygotene), by oocytes and follicular cells (also at a specific  
 CC developmental stage), by haematopoietic cells and blood cells.  
 CC The factors can be used for cell regulation in human hepatocyte,  
 CC haematopoietic or stromal cell culture and to restore or maintain  
 CC cell function in culture. See also AAR34137-R34139.  
 XX  
 SQ Sequence 7 AA;

Query Match 63.48; Score 26; DB 14; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NPTDD 7  
 Db 2 nptde 6  
 |||||

RESULT 2  
 ID AAY17785  
 XX AAY17785 standard; peptide; 7 AA.  
 AC AAY17785;  
 XX  
 XX 12-AUG-1999 (first entry)  
 DT  
 DE B. thuringiensis crystal protein CryIC mutant peptide fragment.  
 XX  
 KW Bacillus thuringiensis; delta-endotoxin; insecticide; crystal protein;  
 KW lepidopteran insect; CryIC; genetic engineering; mutagenesis; mutant;  
 KW caterpillar; beetle; mosquito; toxic; modification.  
 XX  
 OS Synthetic.  
 OS Bacillus thuringiensis.  
 XX  
 XX US5914318-A.  
 PN  
 XX 22-JUN-1999.  
 PD  
 XX  
 XX 26-NOV-1997; 97US-0980071.  
 PF  
 XX 26-NOV-1997; 97US-0980071.  
 PR  
 XX 27-NOV-1996; 96US-0757536.  
 XX  
 XX (ECOG-) ECOGEN INC.  
 PA  
 XX Baum JA, Gilmer AJ, Mettuss AL;  
 PI WPI; 1999-370510/31.  
 XX N-PSDB; AAX80016.  
 DR  
 XX New modified delta-endotoxin crystal proteins from Bacillus  
 PT thuringiensis are useful in insecticidal compositions  
 XX  
 XX Example 5; Fig 4; 144pp; English.

XX The present invention describes a new composition comprising an isolated  
 CC polypeptide for modified Bacillus thuringiensis crystal proteins (CryIC).  
 CC The polypeptide of the composition is insecticidally-active against  
 CC Lepidopterans. The composition is toxic to an insect cell and comprised  
 CC within an insecticidal formulation can be used as a plant protective  
 CC spray which is toxic to caterpillars, beetles and mosquitoes. The  
 CC polypeptide of the composition may be used to kill an insect through  
 CC ingestion of the composition directly or by ingestion of a plant coated  
 CC with the composition or a transgenic plant that expresses the polypeptide  
 CC composition. The insecticidal proteins produced by B. thuringiensis are  
 CC harmless to plants and other non-targeted organisms but toxic to their  
 CC specific target insect. The polypeptides have improved toxicity so a  
 CC reduced amount of bioinsecticide per unit area of treated crop can be  
 CC used allowing economic and efficient utilization in the field. The  
 CC present sequence represents a B. thuringiensis CryIC crystal protein  
 CC mutant peptide fragment.  
 XX  
 SQ Sequence 7 AA;

Query Match 58.5%; Score 24; DB 20; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPT 5  
 Db 3 phnpt 7  
 |||||

RESULT 3  
 ID AAY01210 standard; peptide; 7 AA.  
 XX AAY01210;  
 AC AAY01210;  
 XX  
 XX 26-MAY-1999 (first entry)  
 DT  
 DE Conserved peptide from tomato for designing ACS primers.  
 XX  
 KW 1-aminocyclopropane-1-carboxylic acid synthase; ACS; EFE; banana;  
 KW ethylene forming enzyme; ethylene biosynthesis; plant; fruit ripening;  
 KW transgenic; enzyme; inhibition; flavour; texture; tomato.  
 XX  
 OS Lycopersicon esculentum.  
 XX  
 XX US5886164-A.  
 PN  
 XX 23-MAR-1999.  
 PD  
 XX 15-APR-1996; 96US-0632598.  
 PF  
 XX 15-APR-1996; 96US-0632598.  
 PR  
 XX (ZENE ) ZENECA LTD.  
 XX  
 XX Bird CR, Fletcher JD;  
 PI WPI; 1999-228611/19.  
 DR  
 XX Novel isolated cDNA molecules ((pASC6) and (pACOS7)) encoding  
 PT 1-aminocyclopropane-1-carboxylic acid synthase (ACS) and an ethylene  
 XX forming enzyme (EFE) - useful for modifying fruit ripening  
 XX characteristics, especially in bananas  
 XX  
 XX Example 1; Columns 5-6; 22pp; English.

XX The invention relates to two isolated cDNA molecules ((pASC6) and  
 CC (pACOS7)) encoding 1-aminocyclopropane-1-carboxylic acid synthase (ACS)  
 CC and an ethylene forming enzyme (EFE), respectively. The clones are  
 CC deposited under the Accession Numbers NCIMB 40813 and NCIMB 40814,  
 CC respectively. pASC6 and pACOS7 may be used to genetically control  
 CC ethylene biosynthesis in plants and hence regulate the ethylene-induced

CC processes involved in fruit ripening (and other ethylene related  
 CC processes). Vectors comprising the cDNA sequences may be used to produce  
 CC transgenic bananas with altered fruit ripening characteristics. The  
 CC orientation of the pASC6 and pCOS7 used in the construct, will determine  
 CC how the ripening process is affected. If the genes have a sense  
 CC orientation, and transcribe mRNA that encodes an active enzyme, the rate  
 CC of ripening will be increased (up-regulation) (however, full-length sense  
 CC constructs can also be used to inhibit enzyme expression by co-  
 CC inhibition). If genes encode antisense mRNA, they will inhibit the  
 CC expression of the genes involved in fruit ripening and hence slow the  
 CC process down (down-regulation). In this manner different spatial and  
 CC temporal patterns of genes expression can be produced. Retardation of the  
 CC rate of ripening will reduce the rate of deterioration of banana fruit  
 CC after harvest. This helps in production of high quality fruit that has  
 CC improved flavour and texture. The present sequence represents a conserved  
 CC peptide from tomato used for designing 3' degenerate primers for PCR  
 CC amplification of ACS gene fragments from banana.

XX Sequence 7 AA;

Query Match 58.5%; Score 24; DB 20; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNP 4  
 ||||  
 Db 1 pnp 4

#### RESULT 4

AA82403  
 ID AAY82403 standard; Peptide: 7 AA.

AC AAY82403;

DT 27-JUN-2000 (first entry)

DE B. thuringiensis crystal gene CryIC mutated peptide SEQ ID NO:26.

CC Bacillus thuringiensis; CryIC; crystal protein; insecticide; insect;  
 CC delta-endotoxin; lepidopteran; modification; genetic engineering;  
 CC resistance; mutant; mutagenesis.

OS Bacillus thuringiensis.

OS Synthetic.

PN US6033874-A.

PD 07-MAR-2000.

PF 18-MAY-1999; 99US-0314093.

PR 26-NOV-1997; 97US-0980071.

PR 27-NOV-1996; 96US-0757536.

PA (ECOG-) ECOGEN INC.

PI Mettus AL, Baum JA, Gilmer AJ;

XX WPI; 2000-255697/22.

CC New mutant Bacillus thuringiensis endotoxin, used for controlling  
 CC lepidopteran pests, has mutated loop region to impart higher  
 CC insecticidal activity -

XX Example 1; Fig 4; 153pp; English.

CC The present invention describes isolated Bacillus thuringiensis CryIC  
 CC delta-endotoxin polypeptides having: (i) at least one amino acid (aa)  
 CC mutation in the loop region between alpha-helices 6 and 7 of domain 1;  
 CC and (ii) better activity against Lepidoptera than the native CryIC.  
 CC The polypeptides, possibly after activation in the digestive tract of

CC insects, kills insect cells by formation of pores and disturbing  
 CC cellular homeostasis. The polypeptides are used to control lepidopteran  
 CC pests on plants, either: (i) applied as a composition; or (ii) expressed  
 CC in plants from heterologous nucleic acid (generating insect-resistant  
 CC plants). They are more active against Lepidoptera than native CryIC.  
 CC AAA08144 to AAA08182, and AAY82396 to AAY82432, represent sequences used  
 CC in the exemplification of the present invention.

XX Sequence 7 AA;

Query Match 58.5%; Score 24; DB 21; Length 7;

Best Local Similarity 80.0%; Pred. No. 4.3e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPT 5  
 | |||  
 Db 3 phnpt 7

#### RESULT 5

AAU05017

ID AAU05017 standard; Peptide: 7 AA.

XX AAU05017;

DT 24-OCT-2001 (first entry)

DE Peptide encoded by PCR primers #8-12 used to clone ACS from banana.

XX 1-aminocyclopropane-1-carboxylic acid synthase; ACS; banana;

KW ethylene biosynthesis; ethylene-forming enzyme; EFE; fruit ripening;

XX fruit storage.

OS Lycopersicon esculentum.

OS Cucurbita sp.

OS Vigna sp.

OS Nicotiana sp.

OS Dianthus sp.

OS Orchideae.

XX US6262346-B1.

XX 17-JUL-2001.

XX 15-JAN-1999; 99US-0231240.

XX 15-APR-1996; 96US-0632598.

XX (ZENE ) ZENECA LTD.

XX Bird CR, Fletcher JD;

XX WPI; 2001-450497/48.

CC Modifying level of ethylene biosynthesis in plant of genus Musa,  
 CC involves inserting into genome of plant a DNA sequence encoding banana  
 CC 1-aminocyclopropane-1-carboxylic acid synthase or ethylene-forming  
 CC enzyme -

XX Example 1; Column 6; 23pp; English.

XX The sequence represents the amino acid sequence encoded by PCR primers  
 CC (AAS09910-AAS09915) used to clone 1-aminocyclopropane-1-carboxylic acid  
 CC synthase (ACS) from banana, used in the method of the invention. The  
 CC method involves modifying the level of ethylene biosynthesis in a plant  
 CC of the genus Musa by inserting into the genome of the plant a DNA  
 CC sequence (I) encoding a banana 1-aminocyclopropane-1-carboxylic acid  
 CC synthase (ACS) or an ethylene-forming enzyme (EFE), where (I) is in sense  
 CC or antisense configuration, and modifies the level of activity of ACS or  
 CC EFE. This retards the rate of ripening in banana fruits which reduces the  
 CC rate of deterioration of banana fruit after harvest. As a result, fruit  
 CC may be harvested when they have reached partial or full ripeness and

CC still have the robustness to withstand handling and transport to reach  
 CC the consumer in good condition. In this way high quality ripe fruit can  
 CC be made available to the consumer with reduced requirement for post-  
 CC harvest treatment. High quality fruit will have improved flavour and  
 CC texture. High quality fruit can be produced consistently over a wide  
 CC harvest period, and such fruit can be held in store for long periods and  
 CC ripened to optimal quality by the supply of exogenous ethylene.  
 XX  
 SQ Sequence 7 AA;

Query Match 58.5%; Score 24; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PSNP 4  
 ||||  
 Db 1 psnp 4

RESULT 6  
 AAU05024  
 ID AAU05024 standard; Peptide; 7 AA.  
 XX  
 AC AAU05024;

DT 24-OCT-2001 (first entry)

XX Peptide encoded by PCR primers #1-7 used to clone EFE from banana.

XX 1-aminocyclopropane-1-carboxylic acid synthase; ACS; banana;  
 KW ethylene biosynthesis; ethylene-forming enzyme; EFE; fruit ripening;  
 KW fruit storage; tomato; squash; peach; melon; avocado; mustard;  
 KW apple; carnation.

XX Lycopersicon esculentum.

OS Cucurbita sp.

OS Prunus sp.

OS Synapsis sp.

OS Dianthus sp.

OS Malus sp.

OS Persea sp.

XX US6262346-B1.

PN 17-JUL-2001.

PD 15-JAN-1999; 99US-0231240.

XX 15-APR-1996; 96US-0632598.

XX (ZENE) ZENECA LTD.

PA Bird CR, Fletcher JD;

XX WPI: 2001-450497/48.

XX N-PSDB; AAS09917, AAS09918, AAS09919, AAS09920, AAS09921, AAS09922,

XX AAS09923, AAS09924.

XX Modifying level of ethylene biosynthesis in plant of genus Musa,

XX involves inserting into genome of plant a DNA sequence encoding banana

XX 1-aminocyclopropane-1-carboxylic acid synthase or ethylene-forming

XX enzyme

XX Example 2; Column 8; 23pp; English.

XX The sequence represents the amino acid sequence encoded by PCR primers

XX (AAS09917-AAS09924) used to clone ethylene-forming enzyme (EFE) from

XX banana, used in the method of the invention. The method involves

XX modifying the level of ethylene biosynthesis in a plant of the genus

XX Musa by inserting into the genome of the plant a DNA sequence (I)

XX encoding a banana 1-aminocyclopropane-1-carboxylic acid synthase (ACS)

XX or an ethylene-forming enzyme (EFE), where (I) is in sense or antisense

CC configuration, and modifies the level of activity of ACS or EFE. This  
 CC retards the rate of ripening in banana fruits which reduces the rate of  
 CC deterioration of banana fruit after harvest. As a result, fruit may be  
 CC harvested when they have reached partial or full ripeness and still have  
 CC the robustness to withstand handling and transport to reach the consumer  
 CC in good condition. In this way high quality ripe fruit can be made  
 CC available to the consumer with reduced requirement for post-harvest  
 CC treatment. High quality fruit will have improved flavour and texture.  
 CC High quality fruit can be produced consistently over a wide harvest  
 CC period, and such fruit can be held in store for long periods and ripened  
 CC to optimal quality by the supply of exogenous ethylene.  
 XX  
 SQ Sequence 7 AA;

Query Match 58.5%; Score 24; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PSNP 4  
 ||||  
 Db 1 psnp 4

RESULT 7  
 AAEO6219  
 ID AAEO6219 standard; peptide; 7 AA.  
 XX  
 AC AAEO6219;

DT 25-SEP-2001 (first entry)

XX Ovine Calpha-s peptide.

XX Ovine; Calpha-s; antisense drug; sperm motility; contraceptive;  
 KW cAMP dependent protein kinase catalytic subunit C alpha; therapy;  
 KW male infertility.

XX Ovis sp.

XX WO200148170-A1.

XX 05-JUL-2001.

XX 21-DEC-2000; 2000WO-NO00441.

XX 23-DEC-1999; 99NO-0006424.

XX (REIN/) REINTON N.

XX (ORST/) ORSTAVIK S.

XX (JAHN/) JAHNSEN T.

XX (SKAL/) SKALHEGG B S.

XX (TASK/) TASKEN K.

XX (HAUG/) HAUGEN T.

XX Reinton N, Orstavik S, Jahnsen T, Skalhegg BS, Tasken K, Haugen T;

XX WPI: 2001-418266/44.

XX Novel cDNA sequence encoding C alpha-S protein which is a new splice

XX variant of C alpha catalytic subunit of cAMP dependent protein kinase,

XX useful for treating male infertility and for developing male

XX contraceptive

XX Example 1; Fig 1; 32pp; English.

XX The invention relates to human Calpha-s proteins and cDNA molecules

XX encoding them. Calpha-s is a new splice variant of C alpha catalytic

XX subunit of cAMP dependent protein kinases. Human Calpha-s cDNA is

XX useful for the preparation of an antisense drug. Calpha-s protein

XX is useful for the preparation of a pharmaceutical, and a medicament

XX for manipulating the motility in sperm, for use as contraceptive,

XX and for treating male infertility. A kit comprising antibodies

CC against Calpha-s protein is useful for diagnosing non-motile sperm.  
 CC The present sequence is ovine Calpha-s peptide.

XX  
 SQ Sequence 7 AA;

Query Match 56.1%; Score 23; DB 22; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 2 SNPTD 6  
 III I  
 Db 3 snpnd 7

RESULT 8

AAW78635  
 ID AAW78635 standard; peptide; 7 AA.

XX  
 AC AAW78635;

XX  
 DT 04-NOV-1998 (first entry)

XX  
 DE SH2 domain binding inhibiting peptide SEQ ID NO:128.

XX  
 KW SH2 domain; binding; inhibition; interaction; site specific;  
 KW signal transduction; protein tyrosine kinase; phosphotyrosine;  
 KW growth factor receptor; oncogene; cellular growth; cell proliferation;  
 KW metabolic control; diabetes; PKR; proto-oncogene; insulin receptor.

XX  
 OS Synthetic.

XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 7 /note= "unspecified"

XX  
 PN US5801149-A.

XX  
 PD 01-SEP-1998.

XX  
 PF 21-MAR-1995; 95US-0408604.

XX  
 PR 21-MAR-1995; 95US-0408604.

XX  
 PR 19-JUN-1991; 91US-0722359.

XX  
 PR 09-OCT-1992; 92US-0959949.

XX  
 PR 08-OCT-1993; 93US-0134558.

XX  
 PA (JOSL-) JOSLIN DIABETES CENT INC.

XX  
 PI Shoelson S;

XX  
 DR WPI; 1998-494822/42.

XX  
 PT Inhibiting site-specific SH2 domain interaction - with peptide  
 PT containing phosphotyrosine or phosphotyrosine mimic

XX  
 PS Disclosure; Column 85; 70pp; English.

XX  
 CC A method has been developed of inhibiting a site-specific interaction  
 CC between a first molecule having an SH2 domain and a second molecule that  
 CC interacts with the SH2 domain. The method comprises contacting the first  
 CC molecule with a 4- to 30-mer peptide containing a sequence of formula:  
 CC R1-Met-R3-Met (1), where R1 = phosphotyrosine (pTyr) or a  
 CC phosphotyrosine analogue having a hydrolysis-resistant phosphorous  
 CC moiety, and R3 = any amino acid. AAW78501 to AAW78523 represent  
 CC specifically claimed examples of the peptides described. The peptides are  
 CC useful for modulating both cellular growth to control unwanted cell  
 CC proliferation in e.g. selected malignancies and for metabolic control in  
 CC e.g. diabetes, by inhibiting signal transduction molecules such as  
 CC protein tyrosine kinases (PTKs) which include growth factor receptors,  
 CC proto-oncogene and oncogene products and the insulin receptor. The  
 CC peptides are also useful for treating and for studying the enzymatic  
 CC mechanisms of PTPase activity and the metabolic and biochemical roles of

CC PTPases. AAW78524 to AAW78702 represent other peptides given in the  
 CC present invention, but which are not specifically claimed.

XX  
 SQ Sequence 7 AA;

Query Match 53.7%; Score 22; DB 19; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SNPT 5  
 IIII  
 Db 3 snpt 6

RESULT 9

AAW95063  
 ID AAW95063 standard; peptide; 7 AA.

XX  
 AC AAW95063;

XX  
 DT 19-MAY-1999 (first entry)

XX  
 DE Urokinase peptide derivative.

XX  
 KW Urokinase; anti-invasive; anti-angiogenic; variant; chemical derivative;  
 KW affinity ligand; pharmaceutical; inhibition; tumour invasion; metastasis;  
 KW peptidomimetic; cell migration; proliferation; angiogenesis; fibrosis;  
 KW atherosclerosis; post-balloon angioplasty vascular restenosis; neointima;  
 KW vascular trauma; vascular graft restenosis; inflammatory; lung; scarring;  
 KW wound healing; psoriasis; venous thrombosis.

XX  
 OS Synthetic.

XX  
 PN WO9905263-A1.

XX  
 PD 04-FEB-1999.

XX  
 PF 24-JUL-1998; 98WO-US15437.

XX  
 PR 25-JUL-1997; 97US-0900327.

XX  
 PA (ANGS-) ANGSTROM PHARM INC.

XX  
 PI Jones TR, Mazar AP;

XX  
 DR WPI; 1999-142921/12.

XX  
 PT New urokinase peptides and derivatives with anti-invasive and  
 PT anti-angiogenic activity - useful for treating diseases or  
 PT conditions associated with undesired cell migration, invasion,  
 PT migration-induced proliferation, or angiogenesis

XX  
 PS Example 4; Page 46; 73pp; English.

XX  
 CC The invention relates to an anti-invasive and anti-angiogenic peptide  
 CC compound. The compound has a sequence corresponding to that shown in  
 CC AAW95057 or a substitution variant, addition variant or other chemical  
 CC derivative of that peptide. The peptide or its variant or a derivative  
 CC can be capped or uncapped. The peptides are useful for in vivo or in  
 CC vitro prognostic and diagnostic applications, e.g. as labelled peptides  
 CC to detect a binding site for the peptide on a surface or in the interior  
 CC of a cell (see AAW95057 for detailed uses of the peptide compound, its  
 CC variants/derivatives). The present sequence represents a urokinase  
 CC peptide derivative tested for inhibition of angiogenesis.

XX  
 SQ Sequence 7 AA;

Query Match 53.7%; Score 22; DB 20; Length 7;  
 Best Local Similarity 42.9%; Pred. No. 4.3e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPDD 7  
 Db 1 psspee 7

## RESULT 10

AAW47462  
 ID AAW47462 standard; peptide; 6 AA.

XX AC AAW47462;  
 XX 05-JUN-1998 (first entry)  
 XX Prothrombin/thrombin peptide ligand.  
 XX Prothrombin; thrombin; ligand; affinity matrix; purification.  
 XX Synthetic.

XX EP816377-A2.

XX 07-JAN-1998.

XX 16-JUN-1997; 97EP-0109744.

XX 28-JUN-1996; 96US-0672805.

XX (FARB ) BAYER CORP.

XX Baumbach GA, Buettner JA, Dadd CA, Hammond DJ;

XX WPI; 1998-054836/06.

XX Peptide(s) that bind to prothrombin and thrombin - useful in

XX affinity chromatography

XX Disclosure; Page 6; 21pp; English.

XX The present peptide is a prothrombin/thrombin ligand, which can be

XX used in an affinity matrix for the purification of prothrombin or

XX thrombin.

XX Sequence 6 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 01-DEC-1995; 95WO-US15683.

XX

PR 02-DEC-1994; 94US-0348718.

XX

PA (UYA ) UNIV YALE.

XX

PI Warren SL;

XX

XX WPI; 1996-277787/28.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 01-DEC-1995; 95WO-US15683.

XX

PR 02-DEC-1994; 94US-0348718.

XX

PA (UYA ) UNIV YALE.

XX

PI Warren SL;

XX

XX WPI; 1996-277787/28.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 01-DEC-1995; 95WO-US15683.

XX

PR 02-DEC-1994; 94US-0348718.

XX

PA (UYA ) UNIV YALE.

XX

PI Warren SL;

XX

XX WPI; 1996-277787/28.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 01-DEC-1995; 95WO-US15683.

XX

PR 02-DEC-1994; 94US-0348718.

XX

PA (UYA ) UNIV YALE.

XX

PI Warren SL;

XX

XX WPI; 1996-277787/28.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



XX Peptide conjugate for delivery of bioactive compounds to cell  
PT nucleus - comprises targeting molecule derived from RNA polymerase  
PT II carboxy-terminus and the bioactive molecule  
XX  
XX Claim 7; Page 46; 51pp; English.  
XX  
CC This heptapeptide is derived from RNA polymerase II carboxy-terminus  
CC (Pol II CTD) and is a variation of the consensus motif in AAW18356. At  
CC least two heptapeptide repeats form a peptide conjugate with a linker  
CC and a bioactive molecule. The peptide conjugate can then be used to  
CC deliver the bioactive compound, such as a protein, peptide, sugar or  
CC nucleic acid sequence (e.g. a ribozyme, external guide sequence for  
CC RNase P, antisense sequence, aptamer, triplex forming oligonucleotide,  
CC nucleoside, nucleotide, gene, cDNA, mRNA or RNA) to the nucleus of a  
CC cell. Phosphorylating the peptide alters its association with certain  
CC molecules in the nucleus, such as proteins having a serine/arginine  
CC motif and small sub-nuclear ribonucleoprotein (Sm snRNP) e.g.  
CC phosphorylated Pol II CTD (COOH terminal domain) derived peptides bind to  
CC nuclear proteins associated with transcription and splicing. Also, for  
CC delivery of molecules which are desired or not desired to be in close  
CC association with RNA, it may be desirable to phosphorylate or leave the  
CC peptide unphosphorylated, respectively.  
XX  
XX Sequence 7 AA;

Query Match 51.2%; Score 21; DB 18; Length 7;  
Best Local Similarity 60.0%; Pred. No. 4.3e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSNPT 5  
Db 3 ptspt 7

RESULT 13  
AA40791  
ID AAY40791 standard; peptide; 7 AA.  
AC AAY40791;

DT 01-DEC-1999 (first entry)

XX Amino acid loop 7 (connects scaffold protein beta strands S1 and S2).  
DE Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;  
KW tumour; chemotherapeutic agent.  
XX

OS Synthetic.

PN EP947582-A1.

XX 06-OCT-1999.

PD 31-MAR-1998; 98EP-0870065.

PF 31-MAR-1998; 98EP-0870065.

PR (INNO-) INNOGENETICS NV.

XX Desmet J, Hufton S, Hoogenboom H, Sablon E;

PI WPI; 1999-542958/46.

DR New scaffold protein, useful for stabilizing antigens used as vaccines

XX  
PT  
PT Disclosure; Page 6; 105pp; English.

PS Sequences AAY40791-Y40799 are examples of amino acid loops that can be  
XX used to connect beta strands S1 (AAY40602) and S2 (AAY40603). S1 and S2  
CC are examples of beta strand peptides which form part of a scaffold  
CC

CC protein. Peptides (AAY40601-Y40609) together form a single-chain scaffold  
CC protein which contains at least 1 disulfide bond, contains less than 10%  
CC alpha helix and contains at least 6 beta-strands. The scaffold protein  
CC is constructed of beta strands S1-S6, and may also include beta strands  
CC A1-A3, or any functionally equivalent derivative of these sequences. The  
CC beta strands form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand  
CC connected to the next by hydrogen bonds, which generate a beta sandwich  
CC architecture. If the additional beta strands A1-A3 are included in the  
CC structure the scaffold is constructed of two beta sheets, with the  
CC structures A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are  
CC connected to each other via amino acid loops, where at least one of the  
CC loops binds to a receptor or antigen. The scaffold protein is used to  
CC stabilize antigens or whole proteins such as receptors, or their  
CC fragments. It may be used to bind two separate molecules. For example,  
CC one surface of the scaffold may be bound to a protein which binds to a  
CC tumour antigen. This will target the complex to tumour cells. Another  
CC surface may be bound to a cytotoxic molecule or an autoimmune antibody  
CC which may then kill the tumour cells. Therefore the scaffold protein may  
CC be used to target chemotherapeutic agents to specific cells. It may also  
CC be used to stabilize individual peptides in a peptide library and may be  
CC used in diagnostic techniques, and to stabilize antigens used as  
CC vaccines.  
XX  
XX Sequence 7 AA;

Query Match 51.2%; Score 21; DB 20; Length 7;  
Best Local Similarity 66.7%; Pred. No. 4.3e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPTD 6  
Db 2 pshtnd 7

RESULT 14  
AAY17818  
ID AAY17818 standard; peptide; 7 AA.  
XX  
AC AAY17818;

DT 12-AUG-1999 (first entry)

XX B. thuringiensis crystal protein CryIC mutant peptide fragment.

DE Bacillus thuringiensis; delta-endotoxin; insecticide; crystal protein;  
KW lepidopteran insect; CryIC; genetic engineering; mutagenesis; mutant;  
KW caterpillar; beetle; mosquito; toxic; modification.  
XX

OS Synthetic.

OS Bacillus thuringiensis.

XX US5914318-A.

PN 22-JUN-1999.

PD 26-NOV-1997; 97US-0980071.

PF 26-NOV-1997; 97US-0980071.

PR 27-NOV-1996; 96US-0757536.

XX (ECOG-) ECOGEN INC.

XX Baum JA, Gilmer AJ, Mettuss'AL;

PI WPI; 1999-370510/31.

XX New modified delta-endotoxin crystal proteins from Bacillus  
PT thuringiensis are useful in insecticidal compositions

PS Disclosure; Fig 4; 144pp; English.

XX The present invention describes a new composition comprising an isolated  
CC

CC polypeptide for modified Bacillus thuringiensis crystal proteins (CryIC).  
CC The polypeptide of the composition is insecticidally-active against  
CC Lepidoptera. The composition is toxic to an insect cell and comprised  
CC within an insecticidal formulation can be used as a plant protective  
CC spray which is toxic to caterpillars, beetles and mosquitoes. The  
CC polypeptide of the composition may be used to kill an insect through  
CC ingestion of the composition directly or by ingestion of a plant coated  
CC with the composition or a transgenic plant that expresses the polypeptide  
CC composition. The insecticidal proteins produced by B. thuringiensis are  
CC harmless to plants and other non-targeted organisms but toxic to their  
CC specific target insect. The polypeptides have improved toxicity so a  
CC reduced amount of bioinsecticide per unit area of treated crop can be  
CC used allowing economic and efficient utilization in the field. The  
CC present sequence represents a B. thuringiensis CryIC crystal protein  
CC mutant peptide fragment.  
XX  
SQ Sequence 7 AA;

Query Match 51.2%; Score 21; DB 20; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSNP 4  
Db 3 pnp 6

## RESULT 15

AAAY17819  
ID AAY17819 standard; peptide; 7 AA.

AC AAY17819;

DT 12-AUG-1999 (first entry)

Peptide SEQ ID NO:28 from US5914318.

Bacillus thuringiensis; delta-endotoxin; insecticide; crystal protein;  
Lepidopteran insect; CryIC; genetic engineering; mutagenesis; mutant;  
caterpillar; beetle; mosquito; toxic; modification.

Synthetic.

US5914318-A.

22-JUN-1999.

26-NOV-1997; 97US-0980071.

26-NOV-1997; 97US-0980071.

27-NOV-1996; 96US-0757536.

(ECOG-) ECOGEN INC.

Baum JA, Gilmer AJ, Mettus AL;

WPI; 1999-370510/31.

New modified delta-endotoxin crystal proteins from Bacillus  
thuringiensis are useful in insecticidal compositions

Disclosure; Column 207; 144pp; English.

The present invention describes a new composition comprising an isolated  
polypeptide for modified Bacillus thuringiensis crystal proteins (CryIC).  
The polypeptide of the composition is insecticidally-active against  
Lepidoptera. The composition is toxic to an insect cell and comprised  
within an insecticidal formulation can be used as a plant protective  
spray which is toxic to caterpillars, beetles and mosquitoes. The  
polypeptide of the composition may be used to kill an insect through  
ingestion of the composition directly or by ingestion of a plant coated  
with the composition or a transgenic plant that expresses the polypeptide

CC composition. The insecticidal proteins produced by B. thuringiensis are  
CC harmless to plants and other non-targeted organisms but toxic to their  
CC specific target insect. The polypeptides have improved toxicity so a  
CC reduced amount of bioinsecticide per unit area of treated crop can be  
CC used allowing economic and efficient utilization in the field. The  
CC present sequence represents a peptide from the present invention.  
XX

SQ Sequence 7 AA;

Query Match 51.2%; Score 21; DB 20; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSNP 4

Db 3 pnp 6

Search completed: February 21, 2002, 17:00:39  
Job time: 112 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:59:07 ; Search time 12.51 seconds  
(without alignments)  
12.592 Million cell updates/sec

Title: US-08-753-851-10

Perfect score: 41

Sequence: 1 PSNPTDD 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 35098

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	63.4	7	2 US-08-204-417A-2	Sequence 2, Appl
2	24	58.5	7	2 US-08-632-598-11	Sequence 11, Appl
3	24	58.5	7	2 US-08-980-071-26	Sequence 26, Appl
4	24	58.5	7	2 US-08-757-536-26	Sequence 26, Appl
5	24	58.5	7	3 US-09-314-093-26	Sequence 26, Appl
6	24	58.5	7	4 US-09-250-848-26	Sequence 26, Appl
7	24	58.5	7	4 US-09-251-885-26	Sequence 26, Appl
8	24	58.5	7	4 US-09-231-240-11	Sequence 11, Appl
9	22	53.7	7	1 US-08-408-604A-128	Sequence 128, App
10	21	51.2	6	2 US-08-672-805-14	Sequence 14, Appl
11	21	51.2	7	2 US-08-204-417A-6	Sequence 6, Appl
12	21	51.2	7	2 US-08-980-071-23	Sequence 23, Appl
13	21	51.2	7	2 US-08-980-071-27	Sequence 27, Appl
14	21	51.2	7	2 US-08-980-071-55	Sequence 55, Appl
15	21	51.2	7	2 US-08-757-536-23	Sequence 23, Appl
16	21	51.2	7	2 US-08-757-536-27	Sequence 27, Appl
17	21	51.2	7	2 US-08-757-536-55	Sequence 55, Appl
18	21	51.2	7	3 US-09-314-093-23	Sequence 23, Appl
19	21	51.2	7	3 US-09-314-093-27	Sequence 27, Appl
20	21	51.2	7	3 US-09-314-093-55	Sequence 55, Appl
21	21	51.2	7	3 US-08-566-190-7	Sequence 7, Appl
22	21	51.2	7	4 US-09-250-848-23	Sequence 23, Appl
23	21	51.2	7	4 US-09-250-848-27	Sequence 27, Appl
24	21	51.2	7	4 US-09-250-848-55	Sequence 55, Appl
25	21	51.2	7	4 US-09-251-885-23	Sequence 23, Appl
26	21	51.2	7	4 US-09-251-885-27	Sequence 27, Appl
27	21	51.2	7	4 US-09-251-885-55	Sequence 55, Appl

28 20 48.8 5 6 5217869-83 Patent No. 5217869  
29 20 48.8 7 3 US-08-566-190-14 Sequence 14, Appl  
30 20 48.8 7 3 US-08-566-190-15 Sequence 15, Appl  
31 20 48.8 7 3 US-08-566-190-18 Sequence 18, Appl  
32 19 46.3 4 1 US-08-329-820-19 Sequence 19, Appl  
33 19 46.3 4 2 US-08-441-871-54 Sequence 54, Appl  
34 19 46.3 5 1 US-08-329-820-20 Sequence 20, Appl  
35 19 46.3 5 1 US-08-329-820-129 Sequence 129, App  
36 19 46.3 5 1 US-08-329-820-130 Sequence 130, App  
37 19 46.3 6 1 US-08-192-243-1 Sequence 1, Appl  
38 19 46.3 6 1 US-08-448-059-1 Sequence 1, Appl  
39 19 46.3 6 1 US-08-280-443-34 Sequence 34, Appl  
40 19 46.3 6 1 US-08-457-459-34 Sequence 34, Appl  
41 19 46.3 6 1 US-08-329-820-21 Sequence 21, Appl  
42 19 46.3 6 1 US-08-329-820-132 Sequence 132, App  
43 19 46.3 6 1 US-08-329-820-194 Sequence 194, App  
44 19 46.3 6 1 US-08-329-820-216 Sequence 216, App  
45 19 46.3 6 1 US-08-555-678-34 Sequence 34, Appl

## ALIGNMENTS

RESULT 1  
US-08-204-417A-2  
; Sequence 2, Application US/08204417A  
; Patent No. 5859192  
; GENERAL INFORMATION:  
; APPLICANT: GUILLOUZO, CHRISTIANE  
; APPLICANT: CORLU, ANNE  
; APPLICANT: KNEIP, BERNARD  
; TITLE OF INVENTION: FACTORS FOR THE CELLULAR FUNCTIONAL  
; TITLE OF INVENTION: REGULATION AND BIOLOGICAL APPLICATIONS THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,417A  
; FILING DATE: 19-APR-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9111389  
; FILING DATE: 16-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B.J.  
; REGISTRATION NUMBER: 36,663  
; REFERENCE/DOCKET NUMBER: 1721-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-204-417A-2

Query Match 63.4%; Score 26; DB 2; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NPTDD 7  
|||||  
Db 2 NPTDE 6

## RESULT 2

US-08-632-598-11  
; Sequence 11, Application US/08632598  
; Patent No. 5886164  
; GENERAL INFORMATION:  
; APPLICANT: BIRD, COLIN R  
; APPLICANT: FLETCHER, JONATHAN D  
; TITLE OF INVENTION: RIPENING-RELATED GENES FROM BANANA  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARB Y AND CUSHMAN  
; STREET: 1100 NEW YORK AVENUE N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/632,598  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, PAUL N  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: 223355/SEE50112/US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 861-3000  
; TELEFAX: 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: LYCOPERISICON ESCULENTUM  
; IMMEDIATE SOURCE:  
; CLONE: PROBE 3' PRIMER  
US-08-632-598-11

Query Match 58.5%; Score 24; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSNP 4  
|||||  
Db 1 PSNP 4

## RESULT 3

US-08-980-071-26  
; Sequence 26, Application US/08980071  
; Patent No. 5914318  
; GENERAL INFORMATION:  
; APPLICANT: Baum, James A.  
; APPLICANT: Gilmer, Amy Jelen  
; APPLICANT: Mettus, Anne-Marie Light  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING  
; TITLE OF INVENTION: LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980.071  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/757,536  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara S.  
; REGISTRATION NUMBER: 33,928  
; REFERENCE/DOCKET NUMBER: MECO:206  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-980-071-26

Query Match 58.5%; Score 24; DB 2; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPT 5  
|||||  
Db 3 PHNPT 7

## RESULT 4

US-08-757-536-26  
; Sequence 26, Application US/08757536  
; Patent No. 5942664  
; GENERAL INFORMATION:  
; APPLICANT: Baum, James A.  
; APPLICANT: Gilmer, Amy Jelen  
; APPLICANT: Mettus, Anne-Marie Light  
; TITLE OF INVENTION: Bacillus thuringiensis CryIc  
; TITLE OF INVENTION: Compositions Toxic to Lepidopteran Insects and Methods for  
; TITLE OF INVENTION: Making CryIc Mutants  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White and Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TX  
; COUNTRY: USA  
; ZIP: 77210-4433  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/757,536  
; FILING DATE: CONCURRENTLY HEREWITH  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara

REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MOBT:023  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-757-536-26

Query Match 58.5%; Score 24; DB 2; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPT 5  
Db 3 PHNPT 7

RESULT 5  
US-09-314-093-26  
Sequence 26, Application US/09314093  
Patent No. 6033874  
GENERAL INFORMATION:  
APPLICANT: Baum, James A.  
APPLICANT: Gilmer, Amy Jelen  
APPLICANT: Mettus, Anne-Marie Light  
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING  
LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/314,093  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/980,071  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MECO:206  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-314-093-26

Query Match 58.5%; Score 24; DB 3; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPT 5  
Db 3 PHNPT 7

RESULT 6  
US-09-250-848-26  
Sequence 26, Application US/09250848  
Patent No. 6153814  
GENERAL INFORMATION:  
APPLICANT: Baum, James A.  
APPLICANT: Gilmer, Amy Jelen  
APPLICANT: Mettus, Anne-Marie Light  
TITLE OF INVENTION: Bacillus thuringiensis CryLC  
Compositions Toxic to Lepidopteran Insects and Methods for  
Making CryLC Mutants  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White and Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/250,848  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/757,536  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MOBT:023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-250-848-26

Query Match 58.5%; Score 24; DB 4; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSNPT 5  
Db 3 PHNPT 7

RESULT 7  
US-09-251-885-26  
Sequence 26, Application US/09251885  
Patent No. 6177615  
GENERAL INFORMATION:  
APPLICANT: Baum, James A.  
APPLICANT: Gilmer, Amy Jelen  
APPLICANT: Mettus, Anne-Marie Light  
TITLE OF INVENTION: Bacillus thuringiensis CryLC  
Compositions Toxic to Lepidopteran Insects and Methods for  
Making CryLC Mutants  
NUMBER OF SEQUENCES: 57

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Arnold, White and Durkee  
;; STREET: P.O. Box 4433  
;; CITY: Houston  
;; STATE: TX  
;; COUNTRY: USA  
;; ZIP: 77210-4433  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/251,885  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/757,536  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kitchell, Barbara  
;; REGISTRATION NUMBER: 33,928  
;; REFERENCE/DOCKET NUMBER: MOBT:023  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (512) 418-3000  
;; TELEFAX: (512) 474-7577  
;; INFORMATION FOR SEQ ID NO: 26:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-09-251-885-26

Query Match 58.5%; Score 24; DB 4; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 1 PSNPT 5  
| | | |  
DB 3 PHNPT 7

RESULT 8  
US-09-231-240-11  
;; Sequence 11, Application US/09231240  
;; Patent No. 6262346  
;; GENERAL INFORMATION:  
;; APPLICANT: BIRD, COLIN R  
;; TITLE OF INVENTION: RIPENING-RELATED GENES FROM BANANA  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: CUSHMAN DARRY AND CUSHMAN  
;; STREET: 1100 NEW YORK AVENUE N.W.  
;; CITY: WASHINGTON  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20005-3918  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/231,240  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/632,598  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:

;; NAME: KOKULIS, PAUL N.  
;; REGISTRATION NUMBER: 16,773  
;; REFERENCE/DOCKET NUMBER: 223355/SEE50112/US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 861-3000  
;; TELEFAX: 822-0944  
;; TELEX: 6714627 CUSH  
;; INFORMATION FOR SEQ ID NO: 11:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; ORIGINAL SOURCE:  
;; ORGANISM: LYCOPERSICON ESCULENTUM  
;; IMMEDIATE SOURCE:  
;; CLONE: PROBE 3' PRIMER  
;; US-09-231-240-11

Query Match 58.5%; Score 24; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 1 PSNP 4  
| | | |  
DB 1 PSNP 4

RESULT 9  
US-08-408-604A-128  
;; Sequence 128, Application US/08408604A  
;; Patent No. 5801149  
;; GENERAL INFORMATION:  
;; APPLICANT: Shoelson, Steven  
;; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES  
;; NUMBER OF SEQUENCES: 211  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street, Suite 510  
;; CITY: Boston  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02109-1875  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/408,604A  
;; FILING DATE: 21-MAR-1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/134,558  
;; FILING DATE: 08-OCT-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/959,949  
;; FILING DATE: 09-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/722,359  
;; FILING DATE: 19-JUNE-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Myers, Louis  
;; REGISTRATION NUMBER: 35,965  
;; REFERENCE/DOCKET NUMBER: JDP-014CP3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)227-7400  
;; TELEFAX: (617)227-5941  
;; INFORMATION FOR SEQ ID NO: 128:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: amino acid

;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
US-08-408-604A-128

Query Match 53.7%; Score 22; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SNPT 5  
|||  
Db 3 SNPT 6

RESULT 10  
US-08-672-805-14  
; Sequence 14, Application US/08672805  
; Patent No. 5831003  
; GENERAL INFORMATION:  
; APPLICANT: Baumbach, George A.,  
; APPLICANT: Buettner, Joseph A.,  
; APPLICANT: Dadd, Christopher A.,  
; APPLICANT: Hammond, David J.  
; TITLE OF INVENTION: Peptides Which Bind to Prothrombin and  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bayer Corporation  
; STREET: 800 Dwight Way  
; CITY: Berkeley  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94701-1986  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb Storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,805  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gibling, James A.  
; REGISTRATION NUMBER: 25772  
; REFERENCE/DOCKET NUMBER: MSB-7236  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510)705-7910  
; TELEFAX: (510)705-7904  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6  
; TYPE: amino acid  
; STRANDEDNESS: single strand  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: peptide  
US-08-672-805-14

Query Match 51.2%; Score 21; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.6e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSNP 4  
|:|  
Db 3 PAMP 6

RESULT 11  
US-08-204-417A-6  
; Sequence 6, Application US/08204417A  
; Patent No. 5859192  
; GENERAL INFORMATION:  
; APPLICANT: GUILLOUZO, CHRISTIANE  
; APPLICANT: CORLU, ANNE  
; APPLICANT: KNEIP, BERNARD  
; TITLE OF INVENTION: FACTORS FOR THE CELLULAR FUNCTIONAL  
; TITLE OF INVENTION: REGULATION AND BIOLOGICAL APPLICATIONS THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,417A  
; FILING DATE: 19-APR-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9111389  
; FILING DATE: 16-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B.J.  
; REGISTRATION NUMBER: 36,663  
; REFERENCE/DOCKET NUMBER: 1721-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-204-417A-6

Query Match 51.2%; Score 21; DB 2; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.6e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NPTDD 7  
|||  
Db 2 NPXDE 6

RESULT 12  
US-08-980-071-23  
; Sequence 23, Application US/089800071  
; Patent No. 5914318  
; GENERAL INFORMATION:  
; APPLICANT: Baum, James A.  
; APPLICANT: Gilmer, Amy Jelen  
; APPLICANT: Mettus, Anne-Marie Light  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING  
; TITLE OF INVENTION: LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas



; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,071  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/757,536  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara S.  
; REGISTRATION NUMBER: 33,928  
; REFERENCE/DOCKET NUMBER: MECO:206  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-980-071-23

Query Match 51.2%; Score 21; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.6e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0;

QY 1 PSNP 4  
Db 3 PNNP 6

RESULT 13  
US-08-980-071-27  
; Sequence 27, Application US/08980071  
; Patent No. 5914318  
; GENERAL INFORMATION:  
; APPLICANT: Baum, James A.  
; APPLICANT: Gilmer, Amy Jelen  
; APPLICANT: Mettus, Anne-Marie Light  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,071  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/757,536  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara S.  
; REGISTRATION NUMBER: 33,928  
; REFERENCE/DOCKET NUMBER: MECO:206

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-980-071-27

Query Match 51.2%; Score 21; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.6e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0;

QY 1 PSNP 4  
Db 3 PNNP 6

RESULT 14  
US-08-980-071-55  
; Sequence 55, Application US/08980071  
; Patent No. 5914318  
; GENERAL INFORMATION:  
; APPLICANT: Baum, James A.  
; APPLICANT: Gilmer, Amy Jelen  
; APPLICANT: Mettus, Anne-Marie Light  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,071  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/757,536  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara S.  
; REGISTRATION NUMBER: 33,928  
; REFERENCE/DOCKET NUMBER: MECO:206  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-08-980-071-55

Query Match 51.2%; Score 21; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.6e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0;

QY 1 PSNP 4  
Db 3 PNNP 6

Db 3 PNNP 6

```
RESULT 15
US-08-757-536-23
; Sequence 23, Application US/08757536
; Patent No. 5942664
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Gilmer, Amy Jelen
; APPLICANT: Mettus, Anne-Marie Light
; TITLE OF INVENTION: Bacillus thuringiensis CryIC
; TITLE OF INVENTION: Compositions Toxic to Lepidopteran Insects and Methods for
; TITLE OF INVENTION: Making CryIC Mutants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White and Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,536
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MOBT:023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-757-536-23
```

```
Query Match 51.2%; Score 21; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.6e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 PSNP 4  
Db 3 PNNP 6

Search completed: February 21, 2002, 17:00:58  
Job time: 111 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:03:19 ; Search time 12.74 Seconds  
(without alignments)  
59.792 Million cell updates/sec

Title: US-08-753-851-11

Perfect score: 56

Sequence: 1 TSGGYIFVTF 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 1098

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	24	42.9	10	2	PT0215
2	19	33.9	10	2	PT0243
3	19	33.9	10	2	PH0944
4	18	32.1	9	2	G85802
5	18	32.1	9	2	C57444
6	18	32.1	9	2	PD0443
7	17	30.4	5	2	B61445
8	17	30.4	5	2	C53284
9	17	30.4	5	2	PT0717
10	17	30.4	9	2	PT0080
11	17	30.4	10	2	A60410
12	17	30.4	10	2	H60787
13	17	30.4	10	2	D60788
14	17	30.4	10	2	B60787
15	17	30.4	10	2	A60788
16	17	30.4	10	2	C60589
17	17	30.4	10	2	E60589
18	17	30.4	10	2	B60589
19	17	30.4	10	2	D60588
20	16	28.6	4	2	PT0706
21	16	28.6	5	2	A61445
22	16	28.6	5	2	PT0669
23	16	28.6	5	2	PT0707
24	16	28.6	6	2	JU0355
25	16	28.6	6	2	PT0514
26	16	28.6	6	2	PT0512
27	16	28.6	6	2	PT0727
28	16	28.6	7	1	NYPG7
29	16	28.6	7	2	A60224

30	16	28.6	7	2	PT0542	T-cell receptor be
31	16	28.6	8	2	PI0184	capsid protein VP-
32	16	28.6	8	2	JS0318	leucokinin VIII -
33	16	28.6	8	2	PC4373	telomeric and tetr
34	16	28.6	9	2	I50633	c-rel protein - ch
35	16	28.6	10	2	F60787	sperm-activating p
36	16	28.6	10	2	E39572	sperm-activating p
37	16	28.6	10	2	F60589	sperm-activating p
38	16	28.6	10	2	D60589	sperm-activating p
39	16	28.6	10	2	A60588	lectin GNL1 alpha
40	16	28.6	10	2	S38304	neomycin suppressin -
41	16	28.6	10	2	A56633	cocoonase (EC 3.4.
42	15	26.8	5	2	B61168	neuropeptide Grb-A
43	15	26.8	9	2	A57444	ig heavy chain CDR
44	15	26.8	9	2	B57444	
45	15	26.8	9	2	PT0231	

#### ALIGNMENTS

RESULT 1

PT0215

T-cell receptor beta chain V-J region (4-1-K.1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 30-May-1997

C:Accession: PT0215

R:Nakano, N.; Kikutani, H.; Nishimoto, H.; Kishimoto, T.

J. Exp. Med. 173, 1091-1097, 1991

A:Title: T cell receptor V gene usage of islet beta cell-reactive T cells is not rest

A:Reference number: PT0209; MUID:91217621

A:Accession: PT0215

A:Molecule type: mRNA

A:Residues: 1-10 <NAK>

C:Keywords: T-cell receptor

Query Match 42.9%; Score 24; DB 2; Length 10;

Best Local Similarity 71.4%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SGGYIFY 8

Db 3 SGGYEQY 9

RESULT 2

PT0243

Ig heavy chain CRD3 region (clone 2-103A) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C:Accession: PT0243

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an

A:Reference number: PT0222; MUID:91108337

A:Accession: PT0243

A:Molecule type: DNA

A:Residues: 1-10 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 33.9%; Score 19; DB 2; Length 10;

Best Local Similarity 60.0%; Pred. No. 1.3e+03;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SGGYI 6

Db 6 SGGYL 10

RESULT 3

## PH0944

T-cell receptor beta chain V-D-J region (clone 15) - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
 C:Accession: PH0944  
 R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
 J. Exp. Med. 174, 1467-1476, 1991  
 A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
 A:Reference number: PH0891; MUID:92078857  
 A:Accession: PH0944  
 A:Molecule type: mRNA  
 A:Residues: 1-10 <GOL>  
 A:Experimental source: complete Freund's adjuvant-immunized lymph node  
 A:Note: the authors translated the codon GAC for residue 9 as Glu  
 C:Keywords: T-cell receptor

Query Match 33.9%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 GGY 5  
 Db III  
 6 GGY 8

## RESULT 4

G85802  
 hypothetical protein Z2947 [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: G85802  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: G85802  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-9 <STO>  
 A:Cross-references: GB:AE0051174; NID:g12515957; PIDN:AAG56883.1; GSPDB:GN00145; UWGP:Z29  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z2947

Query Match 32.1%; Score 18; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 YTF 10  
 Db III  
 3 YTF 5

## RESULT 5

C57444  
 neuropeptide Grb-AST B3 - two-spotted cricket  
 C:Species: Gryllus bimaculatus (two-spotted cricket)  
 C:Date: 26-Jan-1996 #sequence\_revision 26-Jan-1996 #text\_change 26-Jan-1996  
 C:Accession: C57444  
 R:Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.  
 J. Biol. Chem. 270, 21103-21108, 1995  
 A:Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cr  
 A:Reference number: A57444; MUID:95403341  
 A:Accession: C57444  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <LOK>

Query Match 32.1%; Score 18; DB 2; Length 9;

Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGGY 5  
 Db III  
 6 SGGW 9

## RESULT 6

PD0443  
 3'-oxoacid CoA-transferase (EC 2.8.3.5) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 05-Feb-1999  
 C:Accession: PD0443  
 R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.  
 submitted to JIPID, August 1998  
 A:Description: Proteome analysis of mouse brain.  
 A:Reference number: PD0441  
 A:Contents: Striatum  
 A:Accession: PD0443  
 A:Molecule type: protein  
 A:Residues: 1-9 <KAW>  
 C:Keywords: CoA-transferase

Query Match 32.1%; Score 18; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FYT 9  
 Db III  
 3 FYT 5

## RESULT 7

B61445  
 Leu-enkephalin - blue mussel  
 C:Species: Mytilus edulis (blue mussel)  
 C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 21-Jan-2000  
 C:Accession: B61445  
 R:Leung, M.K.; Stefano, G.B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 955-958, 1984  
 A:Title: Isolation and identification of enkephalins in pedal ganglia of Mytilus edul  
 A:Reference number: A61445; MUID:84144823  
 A:Accession: B61445  
 A:Molecule type: protein  
 A:Residues: 1-5 <LEU>  
 A:Experimental source: pedal ganglia  
 C:Keywords: neuropeptide; opioid peptide

Query Match 30.4%; Score 17; DB 2; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGYI 6  
 Db III  
 2 GGFL 5

## RESULT 8

C53284  
 T-cell receptor beta 2 chain D region, Dbeta2 - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 02-May-1994 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
 C:Accession: C53284  
 R:Harindranath, N.; Alexander, C.B.; Mage, R.G.  
 Mol. Immunol. 28, 881-888, 1991  
 A:Title: Evolutionarily conserved organization and sequences of germline diversity an  
 A:Reference number: A53284; MUID:91342695  
 A:Accession: C53284  
 A:Status: preliminary  
 A:Molecule type: DNA

A;Residues: 1-5 <HAR>  
 A;Cross-references: GB:S60737; NID:g233916; PIDN:AAB19519, 1; PID:g233919  
 A;Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60740)  
 C;Keywords: T-cell receptor

Query Match 30.4%; Score 17; DB 2; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGG 4  
 |||  
 Db 2 TGGG 5

## RESULT 9

PT0717  
 T-cell receptor beta chain V-D-J region - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C;Accession: PT0717; PT0681

R;Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A;Reference number: PT0509; MUID:91277601

A;Accession: PT0717

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-5 <FEE>

A;Experimental source: newborn thymus, strain BALB/c, 140-2B

A;Accession: PT0681

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-5 <FEZ>

A;Experimental source: day 18 fetal thymus, strain BALB/c, 154-2C

C;Keywords: T-cell receptor

Query Match 30.4%; Score 17; DB 2; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGG 4  
 |||  
 Db 2 SSGG 5

## RESULT 10

PT0080  
 60K Ca binding protein - edible frog (fragment)  
 C;Species: Rana esculenta (edible frog)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
 C;Accession: PT0080  
 R;Treveso, S.; Zorzato, F.; Chiozzi, P.; Melandri, P.; Volpe, P.; Pozzan, T.  
 Biochem. Biophys. Res. Commun. 175, 444-450, 1991  
 A;Title: Frog brain expresses a 60 kDa Ca2+ binding protein similar to mammalian calretinin  
 A;Reference number: PT0080; MUID:91207333

A;Accession: PT0080

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-9 <TRE>

Query Match 30.4%; Score 17; DB 2; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 IFYTF 10  
 :|:|  
 Db 4 VFFKF 8

## RESULT 11

PT0080

## A60410

beta-neoendorphin / dynorphin precursor - guinea pig  
 N;Alternate names: alpha-neoendorphin; proenkephalin B precursor  
 C;Species: Cavia porcellus (guinea pig)  
 C;Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 21-Jan-2000  
 C;Accession: A60410  
 R;Murphy, R.; Turner, C.A.  
 Peptides 11, 65-68, 1990  
 A;Title: Isolation and microsequence analysis of guinea pig alpha-neo-endorphin.  
 A;Reference number: A60410; MUID:90259864  
 A;Accession: A60410  
 A;Molecule type: protein  
 A;Residues: 1-10 <MUR>  
 C;Superfamily: proenkephalin  
 C;Keywords: neuropeptide; opioid peptide

Query Match 30.4%; Score 17; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3e+03;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGYI 6  
 |||  
 Db 2 GGFL 5

## RESULT 12

H60787  
 sperm-activating peptide (Thr-5 speract) - sea urchin (Anthocidaris crassispina)  
 C;Species: Anthocidaris crassispina  
 C;Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 18-Aug-2000  
 C;Accession: H60787  
 R;Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka  
 Comp. Biochem. Physiol. B 89, 687-693, 1988  
 A;Title: Some more speract derivatives associated with eggs of sea urchins, pseudococ  
 A;Reference number: A60787; MUID:88242184  
 A;Accession: H60787

A;Molecule type: protein

A;Residues: 1-10 <SUZ>

C;Comment: This oligopeptide from egg jelly is one of several from this species, all  
 at shows some, but not absolute, species restriction.

C;Superfamily: unassigned animal peptides

Query Match 30.4%; Score 17; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 3e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGG 4  
 |||  
 Db 5 TGGG 8

## RESULT 13

## D60788

sperm-activating peptide (Thr-5 speract) - sea urchin (Pseudocentrotus depressus)  
 C;Species: Pseudocentrotus depressus  
 C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 18-Aug-2000  
 C;Accession: D60788  
 R;Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka  
 Comp. Biochem. Physiol. B 89, 687-693, 1988

A;Title: Some more speract derivatives associated with eggs of sea urchins, pseudococ  
 A;Reference number: A60787; MUID:88242184  
 A;Accession: D60788

A;Molecule type: protein

A;Residues: 1-10 <SUZ>

C;Comment: This oligopeptide from egg jelly is one of several from this species, all  
 at shows some, but not absolute, species restriction.

C;Superfamily: unassigned animal peptides

Query Match 30.4%; Score 17; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 3e+03;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TSGG 4  
| | |  
Db 5 TGGG 8

## RESULT 14

B60787  
sperm-activating peptide (Thr-5 speract) - sea urchin (Hemicentrotus pulcherrimus)  
C:Species: Hemicentrotus pulcherrimus  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 18-Aug-2000  
C:Accession: B60787  
R:Suzuki, N.; Kajiyura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
Comp. Biochem. Physiol. B 89, 687-693, 1988  
A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
A:Reference number: A60787; MUID:88242184  
A:Accession: B60787  
A:Molecule type: protein  
A:Residues: 1-10 <SU2>  
C:Comment: This oligopeptide from egg jelly is one of several from this species, all of  
at shows some, but not absolute, species restriction.  
C:Superfamily: unassigned animal peptides

Query Match 30.4%; Score 17; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 3e+03;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TSGG 4  
| | |  
Db 5 TGGG 8

## RESULT 15

A60788  
sperm-activating peptide (Thr-5 speract) - sea urchin (Strongylocentrotus purpuratus)  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 18-Aug-2000  
C:Accession: A60788  
R:Suzuki, N.; Kajiyura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
Comp. Biochem. Physiol. B 89, 687-693, 1988  
A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
A:Reference number: A60787; MUID:88242184  
A:Accession: A60788  
A:Molecule type: protein  
A:Residues: 1-10 <SU2>  
C:Comment: This oligopeptide from egg jelly is one of several from this species. Unlike  
of the repeats in the known precursor (see PIR:A34543).  
C:Superfamily: unassigned animal peptides

Query Match 30.4%; Score 17; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 3e+03;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TSGG 4  
| | |  
Db 5 TGGG 8

Search completed: February 21, 2002, 17:05:09  
Job time: 110 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:04:54 ; Search time 10.03 Seconds  
(without alignments)  
36.555 Million cell updates/sec

Title: US-08-753-851-11  
Perfect score: 56  
Sequence: 1 TSGGYIFYTF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 334

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	46.4	10	1	ESTA_SCHGA
2	19	33.9	8	1	AL17_CARMA
3	19	33.9	10	1	CU30_LOCOMI
4	17	30.4	8	1	ALL8_CARMA
5	17	30.4	9	1	AL11_CARMA
6	17	30.4	9	1	D1_NEPNO
7	16	28.6	7	1	HY7_PIG
8	16	28.6	8	1	AL15_CARMA
9	16	28.6	8	1	ALL5_CYPDPO
10	16	28.6	8	1	ALL9_CARMA
11	16	28.6	8	1	LCK8_LEUMA
12	16	28.6	9	1	NEUUL_CAVPO
13	16	28.6	10	1	NEMS_DROME
14	15	26.8	5	1	PAP2_PARMA
15	15	26.8	9	1	PGLR_DIAAB
16	15	26.8	10	1	FARP_LOCOMI
17	15	26.8	10	1	LCMS_LEUMA
18	14	25.0	5	1	AL14_CARMA
19	14	25.0	5	1	PSK_DAUCA
20	14	25.0	6	1	FARP_MONEX
21	14	25.0	8	1	AL18_CARMA
22	14	25.0	8	1	ALL3_CYPDPO
23	14	25.0	8	1	ALL4_CALVO
24	14	25.0	8	1	ALL4_CYPDPO
25	14	25.0	9	1	FAR5_ASCSU
26	14	25.0	9	1	FLA2_TREHY
27	14	25.0	9	1	PPK1_PERAM
28	14	25.0	9	1	TKL1_LOCOMI
29	14	25.0	10	1	AL19_CARMA
30	14	25.0	10	1	FARP_MYTED
31	14	25.0	10	1	TKL2_LOCOMI
32	13	23.2	5	1	UC22_MAIZE
33	13	23.2	7	1	ALL2_CARMA

34	13	23.2	7	1	ALL3_CARMA	P81806	carcinus	ma
35	13	23.2	7	1	ALL4_CARMA	P81807	carcinus	ma
36	13	23.2	7	1	ALL5_CARMA	P81808	carcinus	ma
37	13	23.2	8	1	ALL12_CARMA	P81815	carcinus	ma
38	13	23.2	8	1	ALL16_CARMA	P81819	carcinus	ma
39	13	23.2	8	1	ALL1_CYPDPO	P82152	cydia	pomon
40	13	23.2	8	1	ALL6_CYPDPO	P82157	cydia	pomon
41	13	23.2	8	1	ALL7_CARMA	P81809	carcinus	ma
42	13	23.2	8	1	UF06_MOUSE	P38644	mus	musculu
43	13	23.2	8	1	UPAA_HUMAN	P30096	homo	sapien
44	13	23.2	9	1	AL10_CARMA	P81813	carcinus	ma
45	13	23.2	9	1	DSIP_RABIT	P01158	oryctolagus	

## ALIGNMENTS

RESULT 1  
ESTA\_SCHGA  
ID ESTA\_SCHGA STANDARD; PRT; 10 AA.  
AC P81012;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE ESTERASE 52 KDA SUBUNIT (EC 3.1.1.1) (CARBOXYLIC-ESTER HYDROLASE)  
DE (FRAGMENT).  
OS Schizaphis graminum (Aphid).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;  
OC Aphidiformes; Aphidoidea; Aphididae; Aphidini; Schizaphis.  
OX NCBI\_TaxID=13262;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=97468499; PubMed=9327586;  
RA Siegfried B.D., Ono M., Swanson J.J.;  
RT "Purification and characterization of a carboxylesterase associated with organophosphate resistance in the greenbug, Schizaphis graminum (Homoptera: Aphididae).";  
RL Arch. Insect Biochem. Physiol. 36:229-240(1997).  
CC -1- CATALYTIC ACTIVITY: A CARBOXYLIC ESTER + H(2)O = AN ALCOHOL + A CARBOXYLIC ANION.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.  
DR InterPro; IPR002018; Carboxylesterase\_B.  
DR PROSITE; PS00122; CARBOXYLESTERASE\_B.1; PARTIAL.  
DR PROSITE; PS00941; CARBOXYLESTERASE\_B.2; PARTIAL.  
KW Hydrolase; Serine esterase.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1025 MW; 018ABE587865A2C0 CRC64;

Query Match 46.4%; Score 26; DB 1; Length 10;  
Best Local Similarity 83.3%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 SGGYIF 7  
|||||  
Db 5 SGGYDF 10

RESULT 2  
AL17\_CARMA  
ID AL17\_CARMA STANDARD; PRT; 8 AA.  
AC P81820;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CARCINUSTATIN 17.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;

```

RN RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RT Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 33.9%; Score 19; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SGGYIF 7
DB 1 SGGYSF 6

RESULT 3
CU30_LOCM1 STANDARD; PRT; 10 AA.
AC P11735;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CUTICLE PROTEIN 30 (LM-ACP 30) (FRAGMENT).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RX MEDLINE=86108304; PubMed=3943519;
RA Hoejrup P., Andersen S.O., Roepstorff P.;
RT "Isolation, characterization, and N-terminal sequence studies of
cuticular proteins from the migratory locust, Locusta migratoria.";
RL Eur. J. Biochem. 154:153-159(1986).
CC -!- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
CC PIR: C24802; C24802.
DR Structural protein; Cuticle.
KW NON_TER 10
FT MOD_RES 10
SQ SEQUENCE 10 AA; 969 MW; 4973E36B58772877 CRC64;

Query Match 33.9%; Score 19; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGY 5
DB 8 GGY 10

RESULT 4
ALL8_CARMA STANDARD; PRT; 8 AA.
AC P81811;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CARCINUS MAENAS 8.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

```

```

OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RT Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 795 MW; 922879CDCB47687D CRC64;

```

```

Query Match 30.4%; Score 17; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 SGGYIF 7
DB 1 AGPYAF 6

```

```

RESULT 5
ALL1_CARMA STANDARD; PRT; 9 AA.
ID ALL1_CARMA
AC P81814;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CARCINUS MAENAS 11.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RT Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 9
SQ SEQUENCE 9 AA; 927 MW; 832D79CDCB46D861 CRC64;

```

```

Query Match 30.4%; Score 17; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 SGGYIF 7
DB 2 TQGYAF 7

```

```

RESULT 6
DL_NEPNO STANDARD; PRT; 9 AA.
ID DL_NEPNO
AC P24816;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE GASTRIN/CHOLECYSTOKININ-LIKE PEPTIDE DL.

```



OS Nephrops norvegicus (Norway lobster).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;  
 OC Nephropoidea; Nephropidae; Nephrops.  
 OX NCBI\_TaxID=6829;  
 [1]  
 RP SEQUENCE.  
 RN TISSUE=Stomach; PubMed=1747388;  
 RX MEDLINE=92082847; PubMed=1747388;  
 RA Favrel P., Kegel G., Sedimeier D., Keller R., van Wormhoudt A.;  
 RT "Structure and biological activity of crustacean gastrointestinal  
 RT peptides identified with antibodies to gastrin/cholecystokinin.";  
 RL Biochimie 73:1233-1239(1991).  
 CC -1- FUNCTION: MAY CONTROL DIGESTION PROCESSES IN CRUSTACEANS.  
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 DR PIR; A48398; A48398.  
 KW Hormone.  
 SQ SEQUENCE 9 AA; 1038 MW; 60EC79CAB6D8787B CRC64;

Query Match 30.4%; Score 17; DB 1; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1e+05;  
 Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFY 8  
 :|||:  
 DB 1 SEGQDFW 8

RESULT 7  
 HY7\_PIG STANDARD; PRT; 7 AA.  
 ID HY7\_PIG  
 AC P01153;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DE HYPOTHALAMIC HEPTAPEPTIDE.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RX MEDLINE=81213980; PubMed=6263778;  
 RA Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,  
 RA Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;  
 RT "Isolation, structure and synthesis of a heptapeptide with in vitro  
 RT ACTH-releasing activity from porcine hypothalamus.";  
 RL Horm. Metab. Res. 13:228-232(1981).  
 DR PIR; A01417; NYPG7.  
 SQ SEQUENCE 7 AA; 957 MW; 632B45B1FB5059A0 CRC64;

Query Match 28.6%; Score 16; DB 1; Length 7;  
 Best Local Similarity 16.7%; Pred. No. 1e+05;  
 Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 5 YIFYTF 10  
 :|:::  
 DB 1 FIYHSY 6

RESULT 8  
 ALL5\_CARMA STANDARD; PRT; 8 AA.  
 ID ALL5\_CARMA  
 AC P81818;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 15.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunoidea; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 8  
 SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 28.6%; Score 16; DB 1; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SGGYIF 7  
 :|||:  
 DB 1 AGPYSF 6

RESULT 9  
 ALL5\_CVDPO STANDARD; PRT; 8 AA.  
 ID ALL5\_CVDPO  
 AC P82156;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYDIASATIN 5.  
 OS Cydia pomonella (Codling moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Ditrysia;  
 OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 OX NCBI\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
 RA Davey M., East P.D., Thorpe A.;  
 RT "Lepidopteran peptides of the allatostatin superfamily.";  
 RL Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8  
 SQ SEQUENCE 8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match 28.6%; Score 16; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYIF 7  
 :|||:  
 DB 3 GYDF 6

RESULT 10  
 ALL9\_CARMA STANDARD; PRT; 8 AA.  
 ID ALL9\_CARMA  
 AC P81812;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 9.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;

[1]

RP SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;

RX MEDLINE=98121193; PubMed=9461295;

RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,

RA Thorpe A.;

RT "Isolation and identification of multiple neuropeptides of the

RT allatostatin superfamily in the shore crab *Carcinus maenas*."

RL Eur. J. Biochem. 250:727-734(1997).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.

CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

KW Neuropeptide; Amidation; Multigene family.

FT MOD\_RES 8 8

SQ SEQUENCE 8 AA; 781 MW; 7C2879CDBC476878 CRC64;

Query Match 28.6%; Score 16; DB 1; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 1e+05;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GGYIF 7

Db 1 1 1

2 GPYAF 6

RESULT 11

LCK8\_LEUMA

ID LCK8\_LEUMA STANDARD; PRT; 8 AA.

AC P19990;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)

DE LEUCOKININ VIII (L-VIII).

OS Leucophaea maderae (Madeira cockroach).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;

OC Blaberoidea; Blaberidae; Leucophaea.

OX NCBI\_TaxID=6988;

[1]

RP SEQUENCE.

RC TISSUE=Head;

RA Holman G.M., Cook B.J., Nachman R.J.;

RT "Isolation, primary structure and synthesis of leucokinin VII and

RT VIII: the final members of this new family of cephalomyotropic

peptides isolated from head extracts of *Leucophaea maderae*."

RL Comp. Biochem. Physiol. 88C:31-34(1987).

CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE

CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).

CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.

DR PIR; JS0318.

KW Neuropeptide; Amidation.

FT MOD\_RES 8 8

SQ SEQUENCE 8 AA; 902 MW; 736365AB59CAADD8 CRC64;

Query Match 28.6%; Score 16; DB 1; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 1e+05;

Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIFYTF 10

Db 1 1 1

1 GADFYSW 7

RESULT 12

NEUU\_CAVPO

ID NEUU\_CAVPO STANDARD; PRT; 9 AA.

AC P34366;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE NEUROMEDIN U-9 (NMU-9).

GN NMU.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

OX NCBI\_TaxID=10141;

[1]

RP SEQUENCE.

RC TISSUE=Small intestine;

RX MEDLINE=90341105; PubMed=2381877;

RA Murphy R., Turner C.A., Furness J.B., Parker L., Giraud A.;

RT "Isolation and microsequence analysis of a novel form of neuromedin U

RT from guinea pig small intestine."

RL Peptides 11:613-617(1990).

CC -!- FUNCTION: STIMULATES UTERINE SMOOTH MUSCLE CONTRACTION AND CAUSES

CC SELECTIVE VASOCONSTRICTION.

CC -!- SIMILARITY: BELONGS TO THE NMU FAMILY.

DR InterPro; IPR001942; NMU.

DR Pfam; PF02070; NMU; 1.

DR PROSITE; PS00967; NMU; 1.

KW Amidation; Hormone.

FT MOD\_RES 9 9

SQ SEQUENCE 9 AA; 1169 MW; IECF177409C729DB CRC64;

Query Match 28.6%; Score 16; DB 1; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 1e+05;

Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIFY 8

Db 1 1 1

1 GYFLF 5

RESULT 13

NEMS\_DROME

ID NEMS\_DROME STANDARD; PRT; 10 AA.

AC P41494;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE NEOMYOSUPPRESSIN (NEB-MS).

GN NEMS.

OS Drosophila melanogaster (Fruit fly), and

OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227, 7385;

[1]

RP SEQUENCE.

RC SPECIES=D.melanogaster;

RX MEDLINE=93002195; PubMed=1390001;

RA Nichols R.;

RT "Isolation and structural characterization of Drosophila

RT TDVHFLRFamide and FMRFamide-containing neural peptides."

RL J. Mol. Neurosci. 3:213-218(1992).

[2]

RP SEQUENCE.

RC SPECIES=S.bullata; TISSUE=Head;

RX MEDLINE=93047886; PubMed=1358537;

RA Fonagy A., Schoofs L., Proost P., Van Damme J., Bueds H., De Loof A.;

RT "Isolation, primary structure and synthesis of neomyosuppressin, a

RT myoinhibiting neuropeptide from the grey fleshfly, *Neobellieria*

RT bullata."

RL Comp. Biochem. Physiol. 102C:239-245(1992).

CC -!- FUNCTION: MYOINHIBITING NEUROPEPTIDE.

DR Flybase; FBgn0013996; Nems.

KW Neuropeptide; Amidation.

FT MOD\_RES 10 10

SQ SEQUENCE 10 AA; 1248 MW; D3C00329D2C1EAB2 CRC64;

Query Match 28.6%; Score 16; DB 1; Length 10;  
 Best Local Similarity 30.0%; Pred. No. 2.9e+03;  
 Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TSGGYIFYTF 10  
 | : : | |  
 Db 1 TDVDHVFLRF 10

# RESULT 14

PAP2\_PARMA STANDARD; PRT; 5 AA.  
 AC P81864;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PARDAXIN II (PXII) (FRAGMENT).  
 OS Pardachirus marmoratus (Red sea mores sole).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 OC Soleiidae; Soleidae; Pardachirus.  
 OX NCBI\_TaxID=31087;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin secretion;  
 RX MEDLINE=87057369; PubMed=3782138;  
 RA Lazarovici P., Primor N., Loew L.M.;  
 RT "Purification and pore-forming activity of two hydrophobic  
 RT polypeptides from the secretion of the Red sea mores sole (Pardachirus  
 RT marmoratus).";  
 RL J. Biol. Chem. 261:16704-16713(1986).  
 CC -!- FUNCTION: EXHIBITS UNUSUAL SHARK REPELLENT AND SURFACTANT  
 CC PROPERTIES. FORMS VOLTAGE-DEPENDENT, ION-PERMEABLE CHANNELS  
 CC IN MEMBRANES. AT HIGH CONCENTRATION CAUSES CELL MEMBRANE LYSIS.  
 CC -!- SUBUNIT: MONOMER. IN AQUEOUS SOLUTION EXISTS AS A TETRAMER.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.  
 KW Toxin.  
 FT NON\_TER 5 5  
 SQ SEQUENCE 5 AA; 614 MW; 7769C9C9C8100000 CRC64;

Query Match 26.8%; Score 15; DB 1; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYTF 7  
 | : |  
 Db 1 GFFP 4

# RESULT 15

PGLR\_DIAAB STANDARD; PRT; 9 AA.  
 AC P81179;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE ENDO-POLYGALACTURONASE (PG) (EC 3.2.1.15) (FRAGMENT).  
 OS Diaprepes abbreviatus (Sugarcane rootstalk borer weevil).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
 OC Cucujiformia; Phytophaga; Curculionidae; Entiminae; Entimini;  
 OC Diaprepes.  
 OX NCBI\_TaxID=13040;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Larval gut;  
 RA Doodatdar H., McCollum T.G., Mayer R.T.;  
 RT "Purification and characterization of an endo-polygalacturonase from  
 RT the gut of West Indies sugarcane rootstalk borer weevil (Diaprepes

RT abbreviatus L.) larvae.";  
 RL Comp. Biochem. Physiol. 118B:861-867(1997).  
 CC -!- CATALYTIC ACTIVITY: RANDOM HYDROLYSIS OF 1,4-ALPHA-D-  
 CC GALACTOSIDURONIC LINKAGES IN PECTATE AND OTHER GALACTURONANS.  
 CC -!- INDUCTION: INHIBITED BY CITRUS PGIP.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 CC 9.4, ITS MW IS: 44.5 KDA.  
 CC -!- SIMILARITY: WEAK TO OTHER POLYGALACTURONASES.  
 KW Hydrolase; Glycosidase; Cell wall.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1041 MW; 1F49087042DB41BB CRC64;

Query Match 26.8%; Score 15; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 YTF 10  
 | : |  
 Db 2 YTY 4

Search completed: February 21, 2002, 17:07:52  
 Job time: 178 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:04:34 ; Search time 22.09 Seconds  
(without alignments)  
66.217 Million cell updates/sec

Title: US-08-753-851-11  
Perfect score: 56  
Sequence: 1 TSGGVIFVTF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 966

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	30.4	7	10	P82445
2	17	30.4	9	4	Q9H3Y3
3	17	30.4	9	10	Q9S8J8
4	17	30.4	10	6	P82923
5	17	30.4	10	8	Q9TG68
6	16	28.6	8	3	P87225
7	16	28.6	9	13	Q92009
8	15	26.8	10	2	Q52837
9	15	26.8	10	2	O51812
10	15	26.8	10	8	O79924
11	15	26.8	10	8	Q9TG80
12	15	26.8	10	8	Q9TG65
13	15	26.8	10	8	Q9TG62
14	15	26.8	10	8	Q9TG59
15	15	26.8	10	8	Q9TG56
16	15	26.8	10	8	Q9TG53
17	15	26.8	10	8	Q9TG50
18	14	25.0	5	13	P82072
19	14	25.0	5	13	P82073

20	14	25.0	9	4	Q15896	Q15896 homo sapien
21	14	25.0	9	6	Q9TRU7	Q9TRU7 bos taurus
22	14	25.0	9	11	Q9QWR0	Q9QWR0 mus musculu
23	14	25.0	10	8	Q9TG47	Q9TG47 ophisaurus
24	14	25.0	10	8	Q9TG32	Q9TG32 ophisaurus
25	14	25.0	10	8	Q9TFV5	Q9TFV5 eublepharus
26	14	25.0	10	10	P81899	P81899 prunus dulc
27	14	25.0	10	13	Q9PRU9	Q9PRU9 sparus aua
28	13	23.2	8	2	O09258	O09258 synchococc
29	13	23.2	8	2	Q9S443	Q9S443 pseudomonas
30	13	23.2	8	2	Q9R5R2	Q9R5R2 shigella dy
31	13	23.2	8	5	Q9TWH6	Q9TWH6 perinereis
32	13	23.2	8	7	Q9S213	Q9S213 oryctolagus
33	13	23.2	9	2	O30790	O30790 erwinia amy
34	13	23.2	9	4	Q9BYF9	Q9BYF9 homo sapien
35	13	23.2	9	6	Q9TRSO	Q9TRSO oryctolagus
36	13	23.2	10	2	Q44693	Q44693 bacillus am
37	13	23.2	10	2	Q9X534	Q9X534 leclercia a
38	13	23.2	10	2	Q9X533	Q9X533 escherichia
39	13	23.2	10	2	Q9R5N3	Q9R5N3 clostridium
40	13	23.2	10	8	P92733	P92733 fejevaria
41	13	23.2	10	8	P92766	P92766 varanus gri
42	13	23.2	10	8	P92774	P92774 xantusia vi
43	13	23.2	10	8	O79888	O79888 basiliscus
44	13	23.2	10	8	O79897	O79897 hoplocercus
45	13	23.2	10	8	O79915	O79915 leirolepis b

## ALIGNMENTS

RESULT 1

P82445  
ID P82445  
AC P82445; PRELIMINARY; PRT; 7 AA.  
DT 01-JUN-2000 (TRENBLrel. 14, Created)  
DE 01-JUN-2000 (TRENBLrel. 14, Last sequence update)  
DE 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE 10 KDA CELL WALL PROTEIN (FRAGMENT).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. PETIT HAVANA;  
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,  
RA Wojtaszek P., Bolwell G.P.;  
RT "Proteomic study of secondary cell wall proteins from transformed tobacco culture";  
RL Planta 0:0-0(2000).  
CC -1- SUBCELLULAR LOCATION: CELL WALL.  
CC -1- TISSUE SPECIFICITY: XYLEM.  
KW Cell wall.  
FT NON\_TER 7  
SQ SEQUENCE 7 AA; 758 MW; 69D2C1E862D1B2A0 CRC64;

Query Match 30.4%; Score 17; DB 10; Length 7;  
Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIF 7  
Db 4 GRVF 7

RESULT 2

Q9H3Y3  
ID Q9H3Y3 PRELIMINARY; PRT; 9 AA.  
AC Q9H3Y3;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE DJ39B11.1 (CONTIGUES AS DJ461P17.1 IN EM:AL031663) (FRAGMENT).  
 GN DJ39B11.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE: FROM N.A.  
 RA Lloyd D.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL121778; CAB76844.1; -  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1036 MW; 2C417B01B412D1B3 CRC64;

Query Match 30.4%; Score 17; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 4.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGG 4  
 I I I  
 Db 6 TEGG 9

RESULT 3  
 Q9S8J8 PRELIMINARY; PRT; 9 AA.  
 AC Q9S8J8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE OXYZATENSIN-BIOACTIVE PEPTIDE.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95102521; PubMed=7804141;  
 RA Takahashi M., Moriguchi S., Yoshikawa M., Sasaki R.;  
 RT "Isolation and characterization of oryzatensin: a novel bioactive  
 RT peptide with ileum-contracting and immunomodulating activities derived  
 RT from rice albumin.";  
 RL Biochem. Mol. Biol. Int. 33:1151-1158(1994).  
 SQ SEQUENCE 9 AA; 1093 MW; 0E8C67377B56877B CRC64;

Query Match 30.4%; Score 17; DB 10; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 4.7e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIF 8  
 I I I  
 Db 1 GYPY 5

RESULT 4  
 P82923 PRELIMINARY; PRT; 10 AA.  
 ID P82923;  
 AC P82923;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE MITOCHONDRIAL 28S RIBOSOMAL PROTEIN S2 (MRP-S2) (FRAGMENT).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Magnolia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]

RP SEQUENCE.  
 RC TISSUE=LIVER;  
 RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;  
 RT "Small subunit of the mammalian mitochondrial ribosome. Identification  
 RT of the full complement ribosomal proteins present.";  
 RL Submitted (DEC-2000) to the SWISS-PROT data bank.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro: IPR001865; Ribosomal\_S2.  
 DR Pfam: PF00318; Ribosomal\_S2; PARTIAL.  
 DR PRINTS; PR00395; RIBOSOMALS2; PARTIAL.  
 DR PROSITE; PS00962; RIBOSOMAL\_S2\_1; PARTIAL.  
 KW Ribosomal protein; Mitochondrion.  
 FT NON\_TER 1  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1246 MW; 6A7A6679C04B476B CRC64;

Query Match 30.4%; Score 17; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YIF 7  
 I I I  
 Db 5 YIF 7

RESULT 5  
 Q9TG68 PRELIMINARY; PRT; 10 AA.  
 ID Q9TG68;  
 AC Q9TG68;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).  
 OS COI.  
 OS Gerrhonotus liocephalus.  
 OS Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scieroglossa; Anguimorpha; Anguillidae;  
 OC Gerrhonotus.  
 OX NCBI\_TaxID=76654;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99343613; PubMed=10413621;  
 RA Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,  
 RA Papenfuss T.J.;  
 RT "Molecular phylogenetics, tRNA evolution, and historical biogeography  
 RT in anguillid lizards and related taxonomic families.";  
 RL Mol. Phylogenet. Evol. 12:250-272(1999).  
 DR EMBL; AF095614; AAD51532.1; -  
 KW Mitochondrion.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1213 MW; 85E80C733640DC1 CRC64;

Query Match 30.4%; Score 17; DB 8; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 6.2e+03;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGVIFVT 9  
 I : : : I I  
 Db 2 TTARWLFT 10

RESULT 6  
 P87225 PRELIMINARY; PRT; 8 AA.  
 ID P87225  
 AC P87225;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE GIN11 PROTEIN (FRAGMENT).

GN GIN11 OR YLL065W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wedler H., Wedler E., Scharfe M., Wambutt R.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z73169; CAA97518.2; -;  
 DR SGD; S0003988; GIN11.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 1019 MW; 4E21A9C449D5B73B CRC64;

Query Match 28.6%; Score 16; DB 3; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 4.7e+05;  
 Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YIEYTF 10  
 I : : I  
 Db 1 YLSFNF 6

RESULT 7  
 ID Q92009 PRELIMINARY; PRT; 9 AA.  
 AC Q92009;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE C-REL PROTEIN (FRAGMENT).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91133738; PubMed=2284104;  
 RA Hannink M., Temin H.M.;  
 RT "Structure and autoregulation of the c-rel promoter.";  
 RL Oncogene 5:1843-1850(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hannink M., Temin H.M.;  
 RL Oncogene 0:0-0(1990).  
 DR EMBL; X56440; CAA39822.1; -;  
 DR EMBL; X56515; CAA39866.1; -;  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 805 MW; DE317DD87865A2CD CRC64;

Query Match 28.6%; Score 16; DB 13; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGG 4  
 I I I  
 Db 4 SGG 6

RESULT 8  
 ID Q52837 PRELIMINARY; PRT; 10 AA.  
 AC Q52837;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

DE FIXX PROTEIN (10 AA) (FRAGMENT).  
 OS Rhizobium leguminosarum.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=384;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=PRE;  
 RX MEDLINE=90136072; PubMed=2693897;  
 RA Roelvink P.W., Hontelez J.G.J., Van Kammen A., van den Bos R.C.;  
 RT "Nucleotide sequence of the regulatory nifA gene of Rhizobium leguminosarum PRE: transcriptional control sites and expression in Escherichia coli.";  
 RL Mol. Microbiol. 3:1441-1447(1989).  
 DR EMBL; X17073; CAA34923.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1037 MW; 24A5593732C879C8 CRC64;

Query Match 26.8%; Score 15; DB 2; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGY 5  
 I I :  
 Db 2 GGF 4

RESULT 9  
 ID Q51812 PRELIMINARY; PRT; 10 AA.  
 AC Q51812;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE DNA HELICASE (FRAGMENT).  
 GN TRAI.  
 OS Escherichia coli.  
 OC Plasmid F.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92009201; PubMed=1916281;  
 RA Cram D.S., Loh S.M., Cheah K.C.C., Skurray R.A.;  
 RT "Sequence and conservation of genes at the distal end of the transfer region on plasmids F and R6-5.";  
 RL Gene 104:85-90(1991).  
 DR EMBL; M38047; AAA98090.1; -;  
 KW Helicase; Plasmid.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1106 MW; 59A2417731A33B13 CRC64;

Query Match 26.8%; Score 15; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGG 4  
 I I I  
 Db 6 TLGG 9

RESULT 10  
 ID O79924 PRELIMINARY; PRT; 10 AA.  
 AC O79924;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).  
 GN COL.

OS Elgaria panamintina.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae; Elgaria.  
 OX NCBI\_TaxID=5222;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97315309; PubMed=9169559;  
 RA Macey J.R., Larson A., Ananjeva N.B., Papenfuss T.J.;  
 RA "Evolutionary shifts in three major structural features of the  
 RT mitochondrial genome among iguanian lizards.";  
 RL J. Mol. Evol. 44:660-674(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,  
 RA Papenfuss T.J.;  
 RA "Molecular Phylogenetics, tRNA Evolution and Historical Biogeography  
 RT in Anigid Lizards and Related Taxonomic Families.";  
 RL Mol. Phylogenet. Evol. 0:0-0(1998).  
 DR EMBL; U82692; AAC62278.1; -;  
 KW Mitochondrion.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1243 MW; 85EE80C7336411B1 CRC64;

Query Match 26.8%; Score 15; DB 8; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9  
 I: :| |  
 Db 2 TTRWLFST 10

## RESULT 11

QYTG80 PRELIMINARY; PRT; 10 AA.  
 AC QYTG80;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).  
 GN COI.  
 OS Ophiodes striatus.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae; Ophiodes.  
 OX NCBI\_TaxID=102189;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99343613; PubMed=10413621;  
 RA Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,  
 RA Papenfuss T.J.;  
 RA "Molecular Phylogenetics, tRNA evolution, and historical biogeography  
 RT in anigid lizards and related taxonomic families.";  
 RL Mol. Phylogenet. Evol. 12:250-272(1999).  
 DR EMBL; AF085610; AAD51520.1; -;  
 KW Mitochondrion.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1243 MW; 85EE80C7336411B1 CRC64;

Query Match 26.8%; Score 15; DB 8; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9  
 I: :| |  
 Db 2 TTRWLFST 10

## RESULT 12

QYTG65

ID QYTG65 PRELIMINARY; PRT; 10 AA.  
 AC QYTG65;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).  
 GN COI.  
 OS Abronja oaxaca.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae; Abronja.  
 OX NCBI\_TaxID=102176;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99343613; PubMed=10413621;  
 RA Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,  
 RA Papenfuss T.J.;  
 RA "Molecular Phylogenetics, tRNA evolution, and historical biogeography  
 RT in anigid lizards and related taxonomic families.";  
 RL Mol. Phylogenet. Evol. 12:250-272(1999).  
 DR EMBL; AF085615; AAD51535.1; -;  
 KW Mitochondrion.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1243 MW; 85EE80C7336411B1 CRC64;

Query Match 26.8%; Score 15; DB 8; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9  
 I: :| |  
 Db 2 TTRWLFST 10

## RESULT 13

QYTG62 PRELIMINARY; PRT; 10 AA.  
 AC QYTG62;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).  
 GN COI.  
 OS Mesaspis moreleti.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae; Mesaspis.  
 OX NCBI\_TaxID=102187;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99343613; PubMed=10413621;  
 RA Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,  
 RA Papenfuss T.J.;  
 RA "Molecular phylogenetics, tRNA evolution, and historical biogeography  
 RT in anigid lizards and related taxonomic families.";  
 RL Mol. Phylogenet. Evol. 12:250-272(1999).  
 DR EMBL; AF085616; AAD51538.1; -;  
 KW Mitochondrion.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1243 MW; 85EE80C7336411B1 CRC64;

Query Match 26.8%; Score 15; DB 8; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9  
 I: :| |  
 Db 2 TTRWLFST 10

## RESULT 14



Search completed: February 21, 2002, 17:07:35  
Job time: 181 sec

```

Q9TG59
ID Q9TG59 PRELIMINARY; PRT; 10 AA.
AC Q9TG59;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Elgaria coerulea.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguinae; Elgaria.
OX NCBI_TaxID=102184;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99343613; PubMed=10413621;
RA Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,
RA Papenfuss T.J.;
RT "Molecular phylogenetics, trna evolution, and historical biogeography
RT in anquid lizards and related taxonomic families.";
RL Mol. Phylogenet. Evol. 12:250-272(1999).
DR EMBL; AF085617; AAD51541.1; -.
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1243 MW; 85EE80C7336411B1 CRC64;

Query Match 26.8%; Score 15; DB 8; Length 10;
Best Local Similarity 33.3%; Pred. No. 1.4e+04;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9
DB 2 TTRWLFST 10

RESULT 15
Q9TG56
ID Q9TG56 PRELIMINARY; PRT; 10 AA.
AC Q9TG56;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Elgaria kingii.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguinae; Elgaria.
OX NCBI_TaxID=102185;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99343613; PubMed=10413621;
RA Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,
RA Papenfuss T.J.;
RT "Molecular phylogenetics, trna evolution, and historical biogeography
RT in anquid lizards and related taxonomic families.";
RL Mol. Phylogenet. Evol. 12:250-272(1999).
DR EMBL; AF085618; AAD51544.1; -.
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1243 MW; 85EE80C7336411B1 CRC64;

Query Match 26.8%; Score 15; DB 8; Length 10;
Best Local Similarity 33.3%; Pred. No. 1.4e+04;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9
DB 2 TTRWLFST 10

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:01:23 ; Search time 23.61 Seconds  
(without alignments)  
31.374 Million cell updates/sec

Title: US-08-753-851-11

Perfect score: 56

Sequence: 1 TSGGIFITF 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 121109

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	48.2	10	22	AA83483
2	26	46.4	10	18	AAW12680
3	26	46.4	10	21	AB37492
4	26	46.4	10	21	AA81791
5	26	46.4	10	22	AA875680
6	25	44.6	5	22	AAU05475
7	25	44.6	7	22	AAU05547
8	25	44.6	8	21	AB11894
9	25	44.6	8	21	AB11896
10	25	44.6	10	18	AAW12681
11	25	44.6	10	20	AAW78436

12	25	44.6	10	21	AA81792	Urokinase-type pla
13	24	42.9	8	15	AA82161	HIV-1 gp120 V3 loo
14	24	42.9	8	16	AA82627	70K autoantigen pe
15	24	42.9	8	17	AAW94602	Human insulin anal
16	24	42.9	8	20	AAV49914	Glucobacter subo
17	24	42.9	8	21	AA830601	Anti-IgE antibody
18	24	42.9	9	20	AAV28528	Beta-1 integrin ce
19	24	42.9	9	20	AAV04649	Bacterial HsdM sub
20	24	42.9	9	21	AA819060	Amino acid sequenc
21	24	42.9	9	21	AA819064	Amino acid sequenc
22	24	42.9	10	18	AAW12683	Human urokinase-ty
23	24	42.9	10	19	AAW44251	Preferred decapept
24	24	42.9	10	19	AAW45526	Peptide used as hy
25	24	42.9	10	21	AA81794	Urokinase-type pla
26	24	42.9	10	22	AA893846	Human complementar
27	24	42.9	10	22	AA893848	Human complementar
28	24	42.9	10	22	AA893850	Human complementar
29	24	42.9	10	22	AA895676	Human complementar
30	24	42.9	10	22	AA895684	Human complementar
31	24	42.9	10	22	AA896676	Human complementar
32	24	42.9	10	22	AA896688	Human complementar
33	24	42.9	10	22	AA896702	Human complementar
34	24	42.9	10	22	AA896704	Human complementar
35	24	42.9	10	22	AA897087	Saccharomyces cere
36	23	41.1	6	4	AA830581	N-terminally disub
37	23	41.1	6	17	AAW08626	Met-enkephalin ana
38	23	41.1	6	17	AA898865	Cyclic peptide use
39	23	41.1	6	20	AAV28119	Factor Xa inhibiti
40	23	41.1	6	20	AAV28052	Factor Xa inhibiti
41	23	41.1	6	21	AA813825	Interleukin-4 inh
42	23	41.1	9	17	AA897227	Brucella antigen t
43	23	41.1	9	20	AAV07314	Collagen assembly
44	23	41.1	9	20	AAV10309	T cell epitope/MHC
45	23	41.1	9	22	AA867562	Protein encoded by

#### ALIGNMENTS

RESULT 1  
AAG83483  
ID AAG83483 standard; Peptide; 10 AA.  
XX AAG83483;  
XX AC  
XX  
DT 11-SEP-2001 (first entry)  
XX  
XX Arabidopsis thaliana peptide ligand #123.  
DE Arabidopsis thaliana peptide ligand #123.  
XX  
KW Plant; peptide pesticide; peptide herbicide; agricultural research.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX  
PN WO200142279-A2.  
XX  
XX 14-JUN-2001.  
PD  
XX  
PF 13-DEC-2000; 2000WO-GB04781.  
XX  
XX 13-DEC-1999; 99GB-0029469.  
PR  
XX (PROT-) PROTEOM LTD.  
PA  
XX  
XX Roberts GW, Heal JR;  
PI  
XX WPI; 2001-381629/40.  
DR  
XX  
XX A set of peptide ligands for agricultural research and development of  
PT therapeutic agents comprise specific complementary peptides to proteins  
PT encoded by genes of plant genomes -  
XX  
XX Example 4; Page 48; 201pp; English.  
XX

CC The present invention relates to a set of peptide ligands consisting of  
 CC specific complementary peptides to proteins encoded by genes of plant  
 CC genomes. The present sequence is one such peptide from Arabidopsis  
 CC thaliana. The peptides of the present invention are useful in an assay to  
 CC identify a peptide, especially a peptide pesticide or herbicide. The  
 CC peptides are also useful for tools for agricultural research and  
 CC development.  
 XX  
 SQ Sequence 10 AA;

Query Match 48.2%; Score 27; DB 22; Length 10;

Best Local Similarity 44.4%; Pred. No. 1.3e+02;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SGGYIFYTF 10

Db :|| :||

1 aggifyyf 9

RESULT 2

AAW12680

ID AAW12680 standard; peptide; 10 AA.

XX AC AAW12680;

XX DT 06-OCT-1997 (first entry)

XX DE Human urokinase-type plasminogen activator receptor antagonist 1.

XX KW antagonist; human; urokinase-type plasminogen activator receptor; huPAR;

XX KW treat; metastasis; angiogenesis; chronic inflammation; Kaposi's sarcoma;

XX KW diabetic retinopathy; rheumatoid arthritis; inhibit; proliferation;

XX KW tumour cell.

XX OS Synthetic.

XX PN WO9705257-A1.

XX PD 13-FEB-1997.

XX PF 19-JUL-1996; 96WO-US12044.

XX PR 31-JUL-1995; 95US-0509208.

XX PA (CHIR ) CHIRON CORP.

XX PI Bray A, Rosenberg S, Spear KL, Valerio R;

XX DR WPI; 1997-145694/13.

XX PT New peptide antagonists of huPAR - useful for treating uPA-mediated

XX PT diseases, e.g. chronic inflammation or metastasis

XX PS Claim 1; Page 21; 28pp; English.

XX CC AAW12680-90 are peptide antagonists of the human urokinase-type

XX CC plasminogen activator receptor (huPAR) and are used to treat disorders

XX CC involving huPA, especially metastasis, inappropriate angiogenesis,

XX CC chronic inflammation, Kaposi's sarcoma, diabetic retinopathy and

XX CC rheumatoid arthritis. They may also inhibit proliferation of

XX CC huPA-activated tumour cells.

XX SQ Sequence 10 AA;

Query Match 46.4%; Score 26; DB 18; Length 10;

Best Local Similarity 57.1%; Pred. No. 1.9e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGYIFYT 9

Db :|| :||

4 gqylwyt 10

RESULT 3

AAW37492

ID AAW37492 standard; peptide; 10 AA.

XX AC AAW37492;

XX DT 26-FEB-2001 (first entry)

XX DE Peptide linker #1 used as a control peptide.

XX KW Human; fibrinogen; tissue targeting; ionising radiation; GPIIb;

XX KW platelet membrane glycoprotein IIb; platelet membrane glycoprotein IIIa;

XX KW IIIa; platelet activation; tumour; cancer.

XX OS Homo sapiens.

XX PN WO2000066182-A1.

XX PD 09-NOV-2000.

XX PF 28-APR-2000; 2000WO-US11485.

XX PR 29-APR-1999; 99US-0302456.

XX PA (UYVA-) UNIV VANDERBILT.

XX PI Hallahan DE;

XX DR WPI; 2000-687452/67.

XX PT Targeting agents, e.g. therapeutic agents and imaging agents, to target

XX PT tissues, e.g. neoplasms, using ionizing radiation -

XX PS Example 10; Page 94; 135pp; English.

XX CC The present invention relates to a method for targeting a tissue in a

XX CC vertebrate for delivery of an active agent. The method comprises

XX CC exposing the tissue to ionising radiation before, after and/or during

XX CC administration of a delivery vehicle comprising the active agent. The

XX CC ionising radiation targets the delivery vehicle to the tissue. The

XX CC carboxyl-terminal segment of the gamma chain of human fibrinogen (see

XX CC AAW37480) can be used as a delivery vehicle in the present invention.

XX CC Human fibrinogen gamma peptide binds to platelet membrane glycoprotein

XX CC IIb (GPIIb) or platelet membrane glycoprotein IIIa (IIIA) following

XX CC platelet activation, and can be used for site-specific localisation to

XX CC irradiated tumours. The methods of the present invention may be used for

XX CC targeting therapeutic agents and imaging agents to target tissues,

XX CC especially neoplasms, e.g. benign intracranial melanomas, arteriovenous

XX CC malformation, angioma, macular degeneration, melanoma, adenocarcinoma,

XX CC malignant glioma, prostatic carcinoma, kidney carcinoma, bladder

XX CC carcinoma, pancreatic carcinoma, thyroid carcinoma, lung carcinoma,

XX CC colon carcinoma, rectal carcinoma, solid tumours, solid tumour

XX CC metastases, angiofibromas, retrolental fibroplasia, haemangiomas and/or

XX CC Kaposi's sarcomas. The present sequence is a peptide linker. This

XX CC sequence was used as a control peptide in assays for determining the

XX CC effects of the carboxyl-terminal segment of the gamma chain of human

XX CC fibrinogen (AAW37480).

XX SQ Sequence 10 AA;

Query Match 46.4%; Score 26; DB 21; Length 10;

Best Local Similarity 57.1%; Pred. No. 1.9e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SGGYIFY 8

Db :|| :||

3 sgsyvvv 9

RESULT 4

```

AAV81791
ID AAV81791 standard; peptide; 10 AA.
AC AAV81791;
XX
XX
XX 07-JUN-2000 (first entry)
DT
DE
DE Urokinase-type plasminogen activator receptor antagonist peptide #1.
XX
XX Antagonist; urokinase-type plasminogen activator receptor; human;
KW hUPA receptor; hUPA-mediated disorder; metastasis; Kaposi's sarcoma;
KW inappropriate angiogenesis; diabetic retinopathy; chronic inflammation;
KW rheumatoid arthritis; therapy.
XX
XX Synthetic.
OS
XX
XX US6030940-A.
PN
PD
PD 29-FEB-2000.
XX
XX
XX 12-FEB-1997; 97US-0800213.
PF
XX
XX 31-JUL-1995; 95US-0509208.
PR
XX 19-JUL-1996; 96WO-US12044.
PR
XX (CHIR ) CHIRON CORP.
XX
XX
XX Bray A, Valerio R, Rosenberg S, Spear KL;
PI
XX WPI; 2000-205214/18.
DR
XX
XX New peptides are urokinase-type plasminogen activator receptor (uPAR)
PT antagonists useful for treating urokinase-type plasminogen
PT activator-mediated diseases such as metastasis and inappropriate
PT angiogenesis -
XX
XX
XX Claim 1; Column 19; 11pp; English.
XX
XX This sequence represents an antagonist of the invention, and targets
CC human urokinase-type plasminogen activator (hUPA) receptor. The
CC antagonists are used for treating hUPA-mediated disorders, such as
CC metastasis, inappropriate angiogenesis e.g. Kaposi's sarcoma and diabetic
CC retinopathy and chronic inflammation e.g. rheumatoid arthritis.
XX
XX
XX Sequence 10 AA;

Query Match 46.4%; Score 26; DB 21; Length 10;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGYFYT 9
DB 4 GQYIWT 10

RESULT 5
AAB75680
ID AAB75680 standard; Peptide; 10 AA.
AC AAB75680;
XX
XX
XX 10-APR-2001 (first entry)
DT
DE
DE HLA class I binding motif in HOM-TES-84 SEQ ID NO:83.
XX
XX Human; cancer associated antigen precursor; cancer associated antigen;
KW seminoma; HLA; human leukocyte antigen; cytostatic; gene therapy;
KW vaccine; cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200100874-A2.
PN

AAU05475
ID AAU05475 standard; Peptide; 5 AA.
XX
XX
XX AAU05475;
AC
XX
XX 24-OCT-2001 (first entry)
DT
DE
DE Synthetic pentapeptide ligand 22.
XX
XX Pentapeptide ligand; drug design; high throughput screening; HTS; MQSAR;
KW Multivariate Quantitative Structure Activity relationships.
XX
XX Synthetic.
OS
XX
XX WO200136980-A2.
PN
XX
XX 25-MAY-2001.
PD
XX
XX 20-NOV-2000; 2000WO-GB04420.
PF
XX
XX 18-NOV-1999; 99GB-0027346.
PR
XX (NELA-) MELACURE THERAPEUTICS AB.
PA (PETT/) PETT C P.
XX
XX Lundstedt T, Andersson P, Wikberg J, Muceniece R, Prusis P;
XX

```

```

XX
XX 04-JAN-2001.
XX
XX
XX 23-JUN-2000; 2000WO-US17207.
PF
XX
XX 30-JUN-1999; 99US-0346498.
PR
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX
XX Sahin U, Tureci O, Pfreundschuh M;
PI
XX WPI; 2001-112465/12.
XX
XX Diagnosing a disorder characterized by expression of a human cancer
PT associated antigen precursor, comprises detecting interaction of an
PT agent with a nucleic acid molecule encoding the antigen precursor -
XX
XX Example 10; Page 63; 126pp; English.
XX
XX The present invention describes a method for diagnosing a disorder
CC characterised by expression of a human cancer associated antigen (CAA)
CC precursor (I) coded by a NA group 1 nucleic acid molecule (N1)
CC comprising contacting the biological sample with an agent (A) that
CC specifically binds to N1, (I) or its fragment, complexed with an human
CC leukocyte antigen (HLA) molecule and determining the interaction between
CC the agent and N1 or (I). (I) has cytostatic activity and can be used in
CC gene therapy and vaccine production. The method can be used for treating
CC a subject with a condition characterised by expression of (I) in cells
CC of a subject. AAB75607 and AAB75608 represent proteins from human cancer
CC associated antigen precursors, and AAB75609 to AAB75802 represent HLA
CC class I binding motifs in human cancer associated antigen precursors
CC given in the exemplification of the present invention.
XX
XX Sequence 10 AA;

Query Match 46.4%; Score 26; DB 22; Length 10;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIFY 8
DB 5 GYIFF 9

RESULT 6
AAU05475
ID AAU05475 standard; Peptide; 5 AA.
XX
XX
XX AAU05475;
AC
XX
XX 24-OCT-2001 (first entry)
DT
DE
DE Synthetic pentapeptide ligand 22.
XX
XX Pentapeptide ligand; drug design; high throughput screening; HTS; MQSAR;
KW Multivariate Quantitative Structure Activity relationships.
XX
XX Synthetic.
OS
XX
XX WO200136980-A2.
PN
XX
XX 25-MAY-2001.
PD
XX
XX 20-NOV-2000; 2000WO-GB04420.
PF
XX
XX 18-NOV-1999; 99GB-0027346.
PR
XX (NELA-) MELACURE THERAPEUTICS AB.
PA (PETT/) PETT C P.
XX
XX Lundstedt T, Andersson P, Wikberg J, Muceniece R, Prusis P;
XX

```

DR WPI; 2001-432565/46.

PT A novel method for identifying the interaction site, binding site or

PT active site in a macromolecule, using of informative combinatorial

PT chemistry, informative peptide libraries and Multivariate Quantitative

PT Structure Activity Relationships -

XX Example 6; Fig 28; 131pp; English.

XX The sequence represents a pentapeptide ligand selected according to

CC 2 (15-10) fractional factorial design +3 cp (or 1 cp 2 random), which

CC were used to model the interaction of a ligand with its target. The

CC invention relates to characterising the interaction between a Ligand Y

CC and a target X by obtaining information (e.g Multivariate Quantitative

CC Structure Activity Relationships, MQSAR) representing one or more

CC physical and/or chemical properties of targets of type X and type Y

CC to produce a model of interaction. The methods of the invention are

CC useful for identifying outliers of type X or outliers

CC of type Y, drug design, design or identification of lead compounds (e.g

CC by high throughput screening, HTS), design of ligands of type Y with

CC improved affinity and/or selectivity for targets of type X, protein

CC engineering, design of DNA or RNA molecules, design of artificial targets

CC of type X/or artificial ligands of type Y, analysis and/or in the

CC engineering of regions and/or parts of targets of type X and/or ligands

CC of type Y, design of organic compound, catalyst, pharmaceutical, drug,

CC macromolecule being capable of binding a molecule, peptidomimetic,

CC protein, enzyme, antibody, molecule, macromolecule, DNA, RNA or a

CC carbohydrate. The methods are also useful for designing new ligands for

CC known targets and/or for new targets.

XX Sequence 5 AA;

SQ

Query Match 44.6%; Score 25; DB 22; Length 5;

Best Local Similarity 80.0%; Pred. No. 4.3e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYIFY 8

DB 1 gytfy 5

RESULT 7

AAU05547

ID AAU05547 standard; Peptide; 7 AA.

AC AAU05547;

XX 24-OCT-2001 (first entry)

DE Synthetic heptapeptide ligand 22.

XX Heptapeptide ligand; drug design; high throughput screening; HTS; MQSAR;

KW Multivariate Quantitative Structure Activity Relationships.

XX Synthetic.

XX WO200136980-A2.

XX 25-MAY-2001.

XX 20-NOV-2000; 2000WO-GB04420.

XX 18-NOV-1999; 99GB-0027346.

PR (MELA-) MELACURE THERAPEUTICS AB.

PA (PETT/) PETT C P.

XX Lundstedt T, Andersson P, Wikberg J, Muceniece R, Prusis P;

XX WPI; 2001-432565/46.

DR A novel method for identifying the interaction site, binding site or

PT

PT active site in a macromolecule, using of informative combinatorial

PT chemistry, informative peptide libraries and Multivariate Quantitative

PT Structure Activity Relationships -

XX Example 8; Fig 30; 131pp; English.

XX The sequence represents a heptapeptide ligand selected according to

CC 2 (21-16) fractional factorial design with additional experiments

CC added from a half fold over +3 cp, which were used to

CC model the interaction of a ligand with its target. The

CC invention relates to characterising the interaction between a Ligand Y

CC and a target X by obtaining information (e.g Multivariate Quantitative

CC Structure Activity Relationships, MQSAR) representing one or more

CC physical and/or chemical properties of targets of type X and type Y

CC to produce a model of interaction. The methods of the invention are

CC useful for identifying outliers of type X or outliers

CC of type Y, drug design, design or identification of lead compounds (e.g

CC by high throughput screening, HTS), design of ligands of type Y with

CC improved affinity and/or selectivity for targets of type X, protein

CC engineering, design of DNA or RNA molecules, design of artificial targets

CC of type X/or artificial ligands of type Y, analysis and/or in the

CC engineering of regions and/or parts of targets of type X and/or ligands

CC of type Y, design of organic compound, catalyst, pharmaceutical, drug,

CC macromolecule being capable of binding a molecule, peptidomimetic,

CC protein, enzyme, antibody, molecule, macromolecule, DNA, RNA or a

CC carbohydrate. The methods are also useful for designing new ligands for

CC known targets and/or for new targets.

XX Sequence 7 AA;

SQ

Query Match 44.6%; Score 25; DB 22; Length 7;

Best Local Similarity 80.0%; Pred. No. 4.3e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYIFY 8

DB 1 gytfy 5

RESULT 8

AAAB11894

ID AAAB11894 standard; peptide; 8 AA.

AC AAAB11894;

XX 14-NOV-2000 (first entry)

DE Cyclic pseudostellarin F.

XX Split intein; peptide cyclisation; trans splicing; peptide library;

KW peptide screening; cyclic ester intermediate; Ssp DnaE C-intein;

XX plasmid PARCP-p; plasmid PARCPD-p; pseudostellarin F; cyclic.

OS Unidentified.

XX Key Location/Qualifiers

FT Modified-site 1..8

FT /note= "The C-terminus condenses with the N-terminus to

FT form a cyclic peptide"

XX WO200036093-A2.

XX 22-JUN-2000.

XX 18-DEC-1999; 99WO-US30162.

XX 18-DEC-1998; 98US-0112723.

PR 07-OCT-1999; 99US-0158077.

XX (PENN-) PENN STATE RES FOUND.

XX Benkovic SJ, Scott CP, Abel-Santos EV;

XX WPI; 2000-431582/37.  
 XX Non-naturally occurring nucleic acid molecule encodes a target peptide  
 PT produced in a host system in a cyclized form or as a splicing  
 PT intermediate of the cyclized form -  
 XX Examples; Page 44; 92pp; English.  
 XX The invention relates to methods of producing cyclic peptides which  
 CC utilise the trans splicing ability of split inteins to catalyse  
 CC cyclisation of a peptide from a precursor peptide. The precursor peptide  
 CC has a target peptide interposed between two portions of a split intein.  
 CC The interaction of the two portions of the split intein creates a  
 CC catalytically active intein, and also forces the target peptide into a  
 CC loop configuration that stabilises the ester isomer of the amino acid at  
 CC the junction between one of the intein portions and the target peptide.  
 CC A heteroatom from the other intein portion then reacts with the ester to  
 CC form a cyclic ester intermediate. The active intein catalyses the  
 CC formation of an aminosuccinimide that liberates a cyclised form of the  
 CC target peptide, which spontaneously rearranges to form the  
 CC thermodynamically favoured backbone cyclic peptide product. The invention  
 CC also discloses nucleic acids encoding the precursor peptide (i.e. the  
 CC target peptide flanked by the two portions of the split intein),  
 CC expression vectors and host cells comprising the nucleic acid, the  
 CC cyclised peptide product or a splicing intermediate thereof (e.g., a  
 CC thioester or a lariat intermediate), methods of generating peptide  
 CC libraries, and methods of screening peptides. The nucleic acid molecules  
 CC of the invention are used for making cyclic peptides and peptide  
 CC splicing intermediates. The methods can be used for screening the cyclic  
 CC peptides, for identifying molecules that bind to the target peptides and  
 CC to detect the presence of a predetermined characteristic in a peptide  
 CC e.g., the ability to specifically bind a target molecule. The present  
 CC sequence represents cyclic pseudostellarin F which was generated  
 CC in the exemplifications of the invention.  
 XX Sequence 8 AA;  
 SQ

Query Match 44.6%; Score 25; DB 21; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SGGYI 6  
 Db 1 sggy1 5

RESULT 9  
 AAB11896  
 ID AAB11896 standard; Protein; 8 AA.  
 XX  
 AC AAB11896;  
 XX  
 XX 14-NOV-2000 (first entry)  
 XX Pseudostellarin F (linear).  
 XX  
 KW Split intein; peptide cyclisation; trans splicing; peptide library;  
 KW peptide screening; cyclic ester intermediate; Ssp DnaE C-intein;  
 KW plasmid pARCP-p; plasmid pARCPD-p; pseudostellarin F.  
 XX Unidentified.  
 OS  
 XX W0200036093-A2.  
 XX  
 XX 22-JUN-2000.  
 PD  
 XX 18-DEC-1999; 99WO-US30162.  
 PF  
 XX 18-DEC-1998; 98US-0112723.  
 PR  
 PR 07-OCT-1999; 99US-0158077.  
 XX

PA (PENN-) PENN STATE RES FOUND.  
 XX  
 PI Benkovic SJ, Scott CP, Abel-Santos EV;  
 XX  
 DR WPI; 2000-431582/37.  
 DR N-PSDB; AAA61909.  
 XX  
 PT Non-naturally occurring nucleic acid molecule encodes a target peptide  
 PT produced in a host system in a cyclized form or as a splicing  
 PT intermediate of the cyclized form -  
 XX Examples; Fig 3e; 92pp; English.  
 XX The invention relates to methods of producing cyclic peptides which  
 CC utilise the trans splicing ability of split inteins to catalyse  
 CC cyclisation of a peptide from a precursor peptide. The precursor peptide  
 CC has a target peptide interposed between two portions of a split intein.  
 CC The interaction of the two portions of the split intein creates a  
 CC catalytically active intein, and also forces the target peptide into a  
 CC loop configuration that stabilises the ester isomer of the amino acid at  
 CC the junction between one of the intein portions and the target peptide.  
 CC A heteroatom from the other intein portion then reacts with the ester to  
 CC form a cyclic ester intermediate. The active intein catalyses the  
 CC formation of an aminosuccinimide that liberates a cyclised form of the  
 CC target peptide, which spontaneously rearranges to form the  
 CC thermodynamically favoured backbone cyclic peptide product. The  
 CC invention also discloses nucleic acids encoding the precursor peptide  
 CC (i.e. the target peptide flanked by the two portions of the split  
 CC intein), expression vectors and host cells comprising the nucleic acid,  
 CC the cyclised peptide product or a splicing intermediate thereof (e.g., a  
 CC thioester or a lariat intermediate), methods of generating peptide  
 CC libraries, and methods of screening peptides. The nucleic acid molecules  
 CC of the invention are used for making cyclic peptides and peptide  
 CC splicing intermediates. The methods can be used for screening the cyclic  
 CC peptides, for identifying molecules that bind to the target peptides and  
 CC to detect the presence of a predetermined characteristic in a peptide  
 CC e.g., the ability to specifically bind a target molecule. The present  
 CC sequence represents pseudostellarin F in its linear form. Plasmids  
 CC pARCP-p and pARCPD-p encode pseudostellarin F. Expression of these  
 CC plasmids generates cyclic pseudostellarin F (AAB11894).  
 XX Sequence 8 AA;  
 SQ

Query Match 44.6%; Score 25; DB 21; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SGGYI 6  
 Db 1 sggy1 5

RESULT 10  
 AAW12681  
 ID AAW12681 standard; peptide; 10 AA.  
 XX  
 AC AAW12681;  
 XX  
 XX 06-OCT-1997 (first entry)  
 DT  
 XX Human urokinase-type plasminogen activator receptor antagonist 2.  
 DE  
 XX  
 KW antagonist; human; urokinase-type plasminogen activator receptor; hUPAR;  
 KW treat; metastasis; angiogenesis; chronic inflammation; Kaposi's sarcoma;  
 KW diabetic retinopathy; rheumatoid arthritis; inhibit; proliferation;  
 KW tumour cell.  
 XX  
 OS Synthetic.  
 XX  
 XX W09705257-A1.  
 PN  
 XX 13-FEB-1997.  
 PD

XX 19-JUL-1996; 96WO-US12044.  
 XX 31-JUL-1995; 95US-0509208.  
 XX (CHIR ) CHIRON CORP.  
 XX Bray A, Rosenberg S, Spear KL, Valerio R;  
 XX WPI; 1997-145694/13.  
 XX  
 XX New peptide antagonists of huPAR - useful for treating uPA-mediated  
 PT diseases, e.g. chronic inflammation or metastasis  
 PT  
 XX Claim 1; Page 21; 28pp; English.  
 XX  
 XX AAW12680-90 are peptide antagonists of the human urokinase-type  
 CC plasminogen activator receptor (huPAR) and are used to treat disorders  
 CC involving huPA, especially metastasis, inappropriate angiogenesis,  
 CC chronic inflammation, Kaposi's sarcoma, diabetic retinopathy and  
 CC rheumatoid arthritis. They may also inhibit proliferation of  
 CC huPA-activated tumour cells.  
 XX  
 XX Sequence 10 AA;  
 SQ

Query Match 44.6%; Score 25; DB 18; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGYIFYT 9  
 I I I I I  
 Db 4 GCYLYWT 10

RESULT 11  
 AAW78436  
 ID AAW78436 standard; peptide; 10 AA.  
 XX  
 XX AAW78436;  
 XX  
 XX 11-MAY-1999 (first entry)  
 XX  
 XX Human ADAMTS-1 peptide fragment.  
 XX  
 XX Metalloproteinase-disintegrin protein with a thrombospondin domain; ARDS;  
 KW ADAMTS-1; drug composition; foodstuff; leukocyte; thrombocyte; hepatitis;  
 KW blood; erythrocyte; inflammatory disease; rheumatoid arthritis; asthma;  
 KW nephritis; Crohn's disease; acute respiratory disease syndrome.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO9855643-A1.  
 PN  
 XX 10-DEC-1998.  
 PD  
 XX 03-JUN-1998; 98WO-JP02449.  
 PF  
 XX 03-JUN-1997; 97JP-0160422.  
 XX  
 XX (KURE ) KUREHA CHEM IND CO LTD.  
 XX  
 XX Hakoaki M, Hirose K, Inoguchi E, Ishida Y, Ishioka K;  
 PI Kuno K, Matsushima K;  
 PI  
 XX WPI; 1999-070277/06.  
 DR  
 XX Human metalloproteinase-disintegrin protein with thrombospondin  
 PT domain - useful as leukocyte and thrombocyte decreasing and  
 PT erythrocyte increasing agent  
 PT  
 XX Example 1; Page 30; 82pp; Japanese.  
 PS  
 XX

CC This sequence represents a peptide fragment of a human  
 CC metalloproteinase-disintegrin protein with a thrombospondin domain  
 CC (ADAMTS-1). The peptide was used to generate a primer (AAx17991) order  
 CC to isolate the coding sequence (AAx17990). The protein may be used in  
 CC drug compositions and foodstuffs, as an agent for decreasing the  
 CC leukocyte and thrombocyte blood count and increasing the erythrocyte  
 CC blood count, e.g. for treatment of inflammatory diseases such as  
 CC rheumatoid arthritis, hepatitis, nephritis, Crohn's disease, asthma and  
 CC ARDS.  
 XX  
 XX Sequence 10 AA;  
 SQ

Query Match 44.6%; Score 25; DB 20; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TSGGYIFYT 9  
 I I I I I  
 Db 2 TCGGGVGYT 10

RESULT 12  
 AAY81792  
 ID AAY81792 standard; peptide; 10 AA.  
 XX  
 XX AAY81792;  
 XX  
 XX 07-JUN-2000 (first entry)  
 DT  
 XX Urokinase-type plasminogen activator receptor antagonist peptide #2.  
 XX  
 XX Antagonist; urokinase-type plasminogen activator receptor; human;  
 KW huPA receptor; huPA-mediated disorder; metastasis; Kaposi's sarcoma;  
 KW inappropriate angiogenesis; diabetic retinopathy; chronic inflammation;  
 KW rheumatoid arthritis; therapy.  
 XX  
 XX Synthetic.  
 OS  
 XX US6030940-A.  
 PN  
 XX 29-FEB-2000.  
 PD  
 XX 12-FEB-1997; 97US-0800213.  
 PF  
 XX 31-JUL-1995; 95US-0509208.  
 PR  
 XX 19-JUL-1996; 96WO-US12044.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA  
 XX Bray A, Valerio R, Rosenberg S, Spear KL;  
 PI  
 XX WPI; 2000-205214/18.  
 DR  
 XX New peptides are urokinase-type plasminogen activator receptor (uPAR)  
 PT antagonists useful for treating urokinase-type plasminogen  
 PT activator-mediated diseases such as metastasis and inappropriate  
 PT angiogenesis -  
 XX  
 XX Claim 1; Column 19; 11pp; English.  
 PS  
 XX This sequence represents an antagonist of the invention, and targets  
 CC human urokinase-type plasminogen activator (huPA) receptor. The  
 CC antagonists are used for treating huPA-mediated disorders, such as  
 CC metastasis, inappropriate angiogenesis e.g. Kaposi's sarcoma and diabetic  
 CC retinopathy and chronic inflammation e.g. rheumatoid arthritis.  
 CC  
 XX Sequence 10 AA;  
 SQ

Query Match 44.6%; Score 25; DB 21; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;



QY 3 GGYFYTF 9  
 DE 1 1:11  
 DB 4 gcylywt 10

RESULT 13  
 AAR62161  
 ID AAR62161 standard; peptide; 8 AA.  
 AC AAR62161;  
 XX  
 XX 03-MAY-1995 (first entry)  
 XX  
 XX HIV-1 gp120 V3 loop neutralising domain.  
 XX  
 XX epitope; autoantibody; immunoinfective cluster virus;  
 KW nuclear protein antigen; systemic rheumatic disorder;  
 KW human immunodeficiency virus; HIV-1; systemic lupus erythematosus;  
 KW mixed connective tissue disease; scleroderma; glycoprotein 120.  
 XX  
 OS Human Immunodeficiency virus type 1.  
 XX  
 XX WO9420141-A.  
 XX  
 XX 15-SEP-1994.  
 XX  
 XX 10-MAR-1994; 94WO-US02631.  
 XX  
 XX 11-MAR-1993; 93US-0029850.  
 XX  
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 XX  
 XX Douvas A, Ehresmann G, Takehana Y;  
 XX WPI; 1994-302689/37.  
 XX  
 XX Methods for treating immunoinfective cluster virus infections -  
 PT utilise antibodies or fragments characteristic of auto antibodies  
 PT produced by patients with rheumatic disorders  
 XX  
 PS Disclosure; Page 61; 106pp; English.  
 XX  
 XX Previous immunological analyses of the V3 loop of HIV-1 (AAR62159)  
 CC have localised the main neutralising domains. The target of more  
 CC than 80% of neutralising antibodies in HIV-1 infected sera from  
 CC AIDS patients has now been found to overly the consensus binding  
 CC sequence and domain A epitopes of the U1 snRNP 70K protein. In  
 CC AIDS, antibody titres are too low to arrest the disease; however,  
 CC the homologous sequences in 70K are immunodominant targets of  
 CC autoantibodies in the systemic rheumatoid disorder of mixed  
 CC connective tissue disease. The titers of such autoantibodies exceed  
 CC 10 power 7. The anti-snRNP autoantibodies will cross-react with  
 CC HIV-1 epitopes and are useful for treating HIV infection.  
 XX  
 SQ Sequence 8 AA;

Query Match 42.9%; Score 24; DB 15; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIFYTF 10  
 DE 1 1:11  
 DB 2 gfafvtf 8

RESULT 14  
 AAR82627  
 ID AAR82627 standard; Protein; 8 AA.  
 XX  
 AC AAR82627;  
 XX

DT 01-MAY-1996 (first entry)  
 XX  
 DE 70K autoantigen peptide immunologically homologous to HIV-1.  
 XX  
 KW 70K autoantigen; U1 small nuclear ribonucleoprotein; snRNP; HIV-1;  
 KW Human immunodeficiency virus type 1; infection; anti-HIV antibody;  
 KW diagnosis; neutralising epitope.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9525124-A1.  
 XX  
 XX 21-SEP-1995.  
 XX  
 XX 13-MAR-1995; 95WO-US03236.  
 XX  
 XX 14-MAR-1994; 94US-0213743.  
 XX  
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 XX  
 XX Douvas A, Ehresmann G, Takehana Y;  
 PI WPI; 1995-336943/43.  
 XX  
 XX 70K auto:antigen immunologically homologous to HIV-1 - useful for  
 PT diagnosis and therapeutic treatment of HIV-1 in a human  
 XX  
 XX Claim 4; Page 55; 66pp; English.  
 XX  
 XX AAR82609-28 are regions of the 70K autoantigen that are immunologically  
 CC homologous to HIV-1 regions with the major neutralising epitopes of  
 CC gp120/41. The 70K autoantigen is part of the U1 small nuclear  
 CC ribonucleo- protein (snRNP) complex and is immunologically homologous to  
 CC the HIV-1 envelope glycoprotein, gp120/41. The 70K autoantigen can be  
 CC used for diagnosing HIV-1 infection in a subject by identifying the  
 CC presence of anti-HIV-1 antibodies that react with it. The 70K  
 CC autoantigen has a relatively hydrophobic amino terminus and a hydrophilic  
 CC carboxy terminus. Three structures within the 70K autoantigen are  
 CC important for antibody recognition. These structures include the A and B  
 CC domains as well as scattered hydrophilic sequences. Domain B encompasses  
 CC a consensus binding sequence of 8 amino acids that is necessary and  
 CC sufficient for binding to U1 RNA.  
 XX  
 SQ Sequence 8 AA;

Query Match 42.9%; Score 24; DB 16; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIFYTF 10  
 DE 1 1:11  
 DB 1 gfafvtf 7

RESULT 15  
 AAW94602  
 ID AAW94602 standard; peptide; 8 AA.  
 XX  
 AC AAW94602;  
 XX  
 XX 28-APR-1999 (first entry)  
 XX  
 XX Human insulin analogue construction synthetic octapeptide #15.  
 DE  
 DE Insulin; human; hyperglycaemia; modified.  
 XX  
 XX Synthetic.  
 XX  
 XX US5514646-A.  
 XX  
 XX 07-MAY-1996.  
 PD  
 XX

```

PF 05-MAY-1993; 93US-0057201.
XX
PR 17-APR-1991; 91US-0686632.
PR 09-FEB-1989; 89US-0308352.
PR 04-AUG-1989; 89US-0388201.
PR 05-MAY-1993; 93US-0057201.
XX
PA (CHAN/) CHANCE R E.
PA (DIMA/) DIMARCHI R D.
PA (FRAN/) FRANK B H.
PA (SHIE/) SHIELDS J E.
XX
PI Chance RE, DiMarchi RD, Frank BH, Shields JE;
XX
DR WPI; 1996-238795/24.
XX
XX New insulin analogues modified at B-chain position 28 - showing
PT reduced self-association and more rapid onset of activity, used for
PT treating hyperglycaemia
XX
XX Example 19; Column 17; 51pp; English.
XX
XX The present invention describes analogues of human insulin modified
CC at position 29 of the B chain, and optionally at other positions.
CC The insulin analogues have modified physico-chemical and
CC pharmacokinetic properties and are useful in the treatment of
CC hyperglycaemia. They are less prone to dimerisation or self-association
CC to higher molecular weight forms and so possess a comparatively more
CC rapid onset of activity while retaining the biological activity of
CC native human insulin. The present sequence represents a synthetic
CC octapeptide used in the preparation of human insulin analogues in an
CC example from the present invention.
XX
XX Sequence 8 AA;

```

```

Query Match 42.9%; Score 24; DB 17; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 FYTF 10
Db 3 fyt 6

```

Search completed: February 21, 2002, 17:04:31  
Job time: 188 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:02:58 ; Search time 12.47 Seconds  
(without alignments)  
18.046 Million cell updates/sec

Title: US-08-753-851-11

Perfect score: 56

Sequence: 1 TSGGYIFYTF 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 63113

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	46.4	10	3	US-08-800-213A-1
2	25	44.6	10	3	US-08-880-963-17
3	25	44.6	10	3	US-08-800-213A-2
4	24	42.9	8	1	US-08-704-170-67
5	24	42.9	8	3	US-08-331-625A-12
6	24	42.9	8	5	PCT-US94-02631-67
7	24	42.9	8	5	PCT-US95-03236-19
8	24	42.9	8	6	5514646-24
9	24	42.9	10	1	US-08-143-365A-11
10	24	42.9	10	3	US-08-800-213A-4
11	23	41.1	9	4	US-08-849-634B-7
12	22	39.3	6	2	US-09-127-574-5
13	22	39.3	7	1	US-08-273-474-35
14	22	39.3	7	4	US-08-935-100-35
15	22	39.3	7	5	PCT-US94-07881-35
16	22	39.3	8	5	PCT-US95-03236-18
17	22	39.3	9	1	US-08-486-013-4
18	22	39.3	9	2	US-08-482-279-4
19	22	39.3	9	2	US-08-342-268-4
20	22	39.3	9	3	US-08-828-712-17
21	22	39.3	9	3	US-09-015-968-4
22	22	39.3	9	4	US-09-063-276-17
23	22	39.3	9	4	US-09-397-386-4
24	22	39.3	10	1	US-08-282-758B-35
25	22	39.3	10	2	US-08-374-652C-14
26	22	39.3	10	3	US-08-800-213A-5
27	22	39.3	10	3	US-08-800-213A-9

28 39.3 10 5 PCT-US91-02942-88 Sequence 88, Appl  
29 37.5 5 1 US-08-257-782-16 Sequence 16, Appl  
30 37.5 5 1 US-08-577-846-16 Sequence 16, Appl  
31 37.5 5 2 US-08-079-144-16 Sequence 16, Appl  
32 37.5 5 4 US-09-411-531A-4 Sequence 4, Appl  
33 37.5 6 5 PCT-US95-03236-28 Sequence 28, Appl  
34 37.5 7 4 US-09-139-802-142 Sequence 142, App  
35 37.5 7 4 US-08-928-213B-138 Sequence 138, App  
36 37.5 8 1 US-08-279-906A-7 Sequence 7, Appl  
37 37.5 8 1 US-08-279-906A-8 Sequence 8, Appl  
38 37.5 8 1 US-08-279-906A-9 Sequence 9, Appl  
39 37.5 8 1 US-08-310-429A-15 Sequence 15, Appl  
40 37.5 8 2 US-08-690-734A-14 Sequence 14, Appl  
41 37.5 8 3 US-08-742-185-14 Sequence 14, Appl  
42 37.5 8 6 5514646-29 Patent No. 5514646  
43 37.5 9 1 US-07-930-649-6 Sequence 6, Appl  
44 37.5 9 1 US-08-102-738-13 Sequence 13, Appl  
45 37.5 9 1 US-08-486-013-5 Sequence 5, Appl

#### ALIGNMENTS

RESULT 1  
US-08-800-213A-1  
; Sequence 1, Application US/08000213A  
; Patent No. 6030940  
; GENERAL INFORMATION:  
; APPLICANT: Steven Rosenberg, Kerry L. Spear, Robert Valerio,  
; APPLICANT: and Andrew Bray  
; TITLE OF INVENTION: Peptide Analog Inhibitors of Urokinase  
; TITLE OF INVENTION: Receptor Activity  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/800,213A  
; FILING DATE: February 12, 1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fujita, Sharon  
; REGISTRATION NUMBER: 38,459  
; REFERENCE/DOCKET NUMBER: 1118.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510 923 2706  
; TELEFAX: 510 655 3542  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-800-213A-1

Query Match 46.4%; Score 26; DB 3; Length 10;  
Best Local Similarity 57.1%; Pred. No. 71;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGYIFVT 9

Db 1 |::||  
4 GYLYWT 10

## RESULT 2

US-08-880-963-17  
; Sequence 17, Application US/08880963  
; Patent No. 6025470  
; GENERAL INFORMATION:  
; APPLICANT: Valmoril, Danila; Cerottini, Jean-Charles  
; TITLE OF INVENTION: Isolated No. 6025470a-and Decapeptides Which Bind To  
; TITLE OF INVENTION: HLA Molecules, And The Uses Thereof.  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; ZIP: 10022

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
; COMPUTER: IBM

OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/880,963  
; FILING DATE: 23-June-1997  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 6025470man D.

REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5483

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-880-963-17

Query Match 44.6%; Score 25; DB 3; Length 9;

Best Local Similarity 60.0%; Pred. No. 1.6e+05;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIFY 8

Db 1::||

4 GFVYF 8

## RESULT 3

US-08-800-213A-2  
; Sequence 2, Application US/08800213A  
; Patent No. 6030940  
; GENERAL INFORMATION:  
; APPLICANT: Steven Rosenberg, Kerry L. Spear, Robert Valerio,  
; APPLICANT: and Andrew Bray  
; TITLE OF INVENTION: Peptide Analog Inhibitors of Urokinase  
; TITLE OF INVENTION: Receptor Activity  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/800,213A

FILING DATE: February 12, 1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Fujita, Sharon

REGISTRATION NUMBER: 38,459

REFERENCE/DOCKET NUMBER: 1118.002

TELEPHONE: 510 923 2706

TELEFAX: 510 655 3542

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-800-213A-2

Query Match 44.6%; Score 25; DB 3; Length 10;

Best Local Similarity 57.1%; Pred. No. 1.1e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGYIFYT 9

Db 1::||

4 GCYLYWT 10

## RESULT 4

US-08-704-170-67

; Sequence 67, Application US/08704170

; Patent No. 5707626

; GENERAL INFORMATION:

APPLICANT: Douvas, Angeline

APPLICANT: Takehana, Yoshi

APPLICANT: Ehresmann, Glenn

TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR

TITLE OF INVENTION: IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Robbins, Berliner & Carson

STREET: 201 No. 5707626th Figueroa Street, Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90012

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/704,170

FILING DATE:

PRIOR APPLICATION: 424

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/029,850

ATTORNEY/AGENT INFORMATION:

NAME: Spitals, John P.

REGISTRATION NUMBER: 29,215

REFERENCE/DOCKET NUMBER: 1920-331

TELEPHONE: (213) 977-1001

TELEFAX: (213) 977-1003

INFORMATION FOR SEQ ID NO: 67:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

;  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-704-170-67

Query Match 42.9%; Score 24; DB 1; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIFYTF 10  
; : : : :  
Db 2 GFAFVTF 8

## RESULT 5

US-08-331-625A-12  
; Sequence 12, Application US/08331625A  
; Patent No. 6057436  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Timothy J.  
; APPLICANT: Klepfer, Sharon  
; APPLICANT: Reed, Albert Paul  
; APPLICANT: Jones, Elaine V.  
; TITLE OF INVENTION: Canine Coronavirus S Gene and Uses  
; TITLE OF INVENTION: Therefor  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation - Corporate  
; ADDRESSES: Patents  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-2799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/331,625A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/880,194  
; FILING DATE: 08-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/698,927  
; FILING DATE: 13-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/613,066  
; FILING DATE: 14-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schreck, Patricia A.  
; REGISTRATION NUMBER: 33,777  
; REFERENCE/DOCKET NUMBER: SBC H85010-1  
; TELEPHONE: (215) 270-5015  
; TELEFAX: (215) 270-5090  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-331-625A-12

Query Match 42.9%; Score 24; DB 3; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.6e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 YIFYTF 10  
; : : : :  
Db 2 FLFHTF 7

## RESULT 6

PCT-US94-02631-67  
; Sequence 67, Application PC/TUS9402631  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; TITLE OF INVENTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 North Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/02631  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/029,850  
; FILING DATE: 11-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-331  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US94-02631-67

Query Match 42.9%; Score 24; DB 5; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYIFYTF 10  
; : : : :  
Db 2 GFAFVTF 8

## RESULT 7

PCT-US95-03236-19  
; Sequence 19, Application PC/TUS9503236  
; GENERAL INFORMATION:  
; APPLICANT: University of Southern California  
; TITLE OF INVENTION: Methods to Diagnose and Treat HIV-1  
; TITLE OF INVENTION: Infection  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA

ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03236  
FILING DATE: 13-MAR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Imbra, Richard J.  
REGISTRATION NUMBER: 37,643  
REFERENCE/DOCKET NUMBER: FP-SI 1394  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-03236-19

Query Match 42.9%; Score 24; DB 5; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 2; Indels 0;

QY 4 GYIFYTF 10  
I: | | | |  
Db 1 GFAFVTF 7

RESULT 8  
5514646-24  
Patent No. 5514646  
APPLICANT: CHANCE, RONALD E.; DIMARCHI, RICHARD D.; FRANK,  
BRUCE H.; SHIELDS, JAMES E.  
TITLE OF INVENTION: INSULIN ANALOGS MODIFIED AT POSITION  
29 OF THE B CHAIN  
NUMBER OF SEQUENCES: 52  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/57,201  
FILING DATE: 05-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 686,632  
FILING DATE: 17-APR-1991  
APPLICATION NUMBER: 388,201  
FILING DATE: 04-AUG-1989  
APPLICATION NUMBER: 308,352  
FILING DATE: 09-FEB-1989  
SEQ ID NO: 24:  
LENGTH: 8  
5514646-24

Query Match 42.9%; Score 24; DB 6; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 7 FYTF 10  
| | | |  
Db 3 FYTF 6

RESULT 9  
US-08-143-365A-11  
Sequence 11, Application US/08143365A  
Patent No. 5726292  
GENERAL INFORMATION:  
APPLICANT: Lowell, George H

TITLE OF INVENTION: IMMUNO-POTENTIATING SYSTEMS FOR  
PREPARATION OF IMMUNOGENIC MATERIALS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hendricks and Assoc.  
STREET: 9669 A Main Street  
CITY: Fairfax  
STATE: VA  
COUNTRY: US  
ZIP: 22031  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/143,365A  
FILING DATE: 29-OCT-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenna M  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: Lowell-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 425-4250  
TELEFAX: (703) 425-2767  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-143-365A-11

Query Match 42.9%; Score 24; DB 1; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGYIF 7  
| | | |  
Db 1 GGYCF 5

RESULT 10  
US-08-800-213A-4  
Sequence 4, Application US/08800213A  
Patent No. 6030940  
GENERAL INFORMATION:  
APPLICANT: Steven Rosenberg, Kerry L. Spear, Robert Valerio,  
and Andrew Bray  
TITLE OF INVENTION: Peptide Analog Inhibitors of Urokinase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/800,213A  
FILING DATE: February 12, 1997

```

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1118.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510 923 2706
; TELEFAX: 510 655 3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 8
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /label= Xaa
; OTHER INFORMATION: /note= "1-naphthylalanine"
US-08-800-213A-4

```

```

Query Match 42.9%; Score 24; DB 3; Length 10;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 GGYFYFT 9
| | | |
DB 4 GQYLXYT 10

```

```

RESULT 11
US-08-849-634B-7
; Sequence 7, Application US/08849634B
; Patent No. 6296855
; GENERAL INFORMATION:
; APPLICANT: INNOGENETICS N.V.
; TITLE OF INVENTION: NEW 17 kDa BRUCELLA ABORTUS ANTIGEN, RECOMBINANT
; TITLE OF INVENTION: POLYPEPTIDES, NUCLEIC ACIDS CODING FOR THE SAME AND USE
; TITLE OF INVENTION: THEREOF IN DIAGNOSTIC AND PROPHYLACTIC METHODS AND KITS
; FILE REFERENCE: INNS006
; CURRENT APPLICATION NUMBER: US/08/849,634B
; CURRENT FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Brucella abortus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (7)..(9)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-08-849-634B-7

```

```

Query Match 41.1%; Score 23; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 GYIF 7
| | | |
DB 2 GYIF 5

```

```

RESULT 12
US-09-127-574-5
; Sequence 5, Application US/09127574
; Patent No. 5985836

```

```

; GENERAL INFORMATION:
; APPLICANT: Bastek, Patrick
; APPLICANT: Lang, John M.
; APPLICANT: Baumbach, George A.
; APPLICANT: Carbonell, Ruben G.
; TITLE OF INVENTION: Alpha-1 Protease Inhibitor Binding Peptides
; FILE REFERENCE: MSB-7248
; CURRENT APPLICATION NUMBER: US/09/127,574
; CURRENT FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-127-574-5

```

```

Query Match 39.3%; Score 22; DB 2; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.6e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 YIFYT 9
| | | |
DB 2 FIYFT 6

```

```

RESULT 13
US-08-273-474-35
; Sequence 35, Application US/08273474
; Patent No. 5691134
; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David R.
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS
; TITLE OF INVENTION: OF DETECTION UTILIZING THE SAME
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Building, 127
; STREET: Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,474
; FILING DATE: 13-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-273-474-35

```

```

Query Match 39.3%; Score 22; DB 1; Length 7;

```

Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 IFYT 9  
||||  
Db 4 IFYT 7

## RESULT 14

US-08-935-100-35  
; Sequence 35, Application US/08935100  
; Patent No. 6143494  
; GENERAL INFORMATION:  
; APPLICANT: Kilpatrick, David R.  
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS  
; TITLE OF INVENTION: OF DETECTION UTILIZING THE SAME  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
; STREET: Suite 1200, The Candler Building, 127  
; STREET: Peachtree Street, NE  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303-1811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/935,100  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/273,474  
; FILING DATE: 13-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perryman, David G.  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.617  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 688-0770  
; TELEFAX: (404) 688-9880  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-935-100-35

Query Match 39.3%; Score 22; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 IFYT 9  
||||  
Db 4 IFYT 7

## RESULT 15

PCT-US94-07881-35  
; Sequence 35, Application PC/TUS9407881  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of  
; APPLICANT: Department of Health and Human Services  
; APPLICANT: National Institutes of Health  
; APPLICANT: Office of Technology Transfer  
; APPLICANT: Bethesda, Maryland 20892  
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND  
America, as re

NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, The Candler Building,  
STREET: 127 Peachtree Street NE  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303-1811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07881  
FILING DATE: 13 July 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/092,110  
FILING DATE: 13 July 1993  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-07881-35

Query Match 39.3%; Score 22; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 IFYT 9  
||||  
Db 4 IFYT 7

Search completed: February 21, 2002, 17:04:50  
Job time: 112 sec



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: February 21, 2002, 17:07:09 ; Search time 12.69 Seconds  
(without alignments)  
186,084 Million cell updates/sec

Title: US-08-753-851-12  
Perfect score: 164  
Sequence: 1 LCLVPLSLAQIDLNITCRFAGVFHVEKNGRY 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 6512

Minimum DB seq length: 0  
Maximum DB seq length: 31

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	32.3	29	2 I77372	CD44SP - human
2	34	20.7	29	2 S78310	hypothetical prote
3	32	19.5	26	2 S00579	hypothetical prote
4	31	18.9	22	2 F44957	osmotin homolog -
5	30	18.3	17	2 A60889	olfactory glycopro
6	30	18.3	28	2 S10052	ribosomal protein
7	29	17.7	18	2 A43826	dnak-type molecula
8	29	17.7	18	2 H64711	hypothetical prote
9	29	17.7	28	2 I45911	dnak-type molecula
10	29	17.7	30	2 B82233	hypothetical prote
11	28	17.1	19	2 S31613	beta-1,3-glucanase
12	28	17.1	23	2 A83698	hypothetical prote
13	28	17.1	31	2 A81078	hypothetical prote
14	28	17.1	31	2 D82348	hypothetical prote
15	27.5	16.8	26	2 B61525	beta-lactoglobulin
16	27	16.5	29	2 A35121	hypothetical prote
17	26.5	16.2	26	2 D61590	beta-lactoglobulin
18	26.5	16.2	31	2 T07337	plastoquinol--plas
19	26	15.9	23	2 S06660	tissue kallikrein
20	26	15.9	23	2 A44626	dihydropyrimidine
21	26	15.9	23	2 S12502	capsid protein - S
22	26	15.9	23	2 B31657	major fecal allerg
23	26	15.9	28	2 I55596	lysosomal acid lip
24	26	15.9	29	2 T12250	NADH dehydrogenase
25	26	15.9	29	2 S46211	kallikrein rK8 (pk
26	26	15.9	31	2 E53480	T-cell receptor al
27	26	15.9	31	2 T08489	hypothetical prote
28	25.5	15.5	21	2 F64121	hypothetical prote
29	25.5	15.5	22	2 S59071	secretory protein

30	25.5	15.5	25	2 A32248	Ig kappa chain V r
31	25.5	15.5	27	2 T13836	NADH dehydrogenase
32	25.5	15.5	28	2 FC1001	Ig light chain V r
33	25.5	15.5	29	2 H31485	Ig kappa chain V r
34	25	15.2	19	2 A61144	probable flagellar
35	25	15.2	20	2 PQ0710	DNA primase (EC 2.
36	25	15.2	21	2 D48518	PEP1 5'-region hyp
37	25	15.2	25	2 A25865	gamma-crystallin I
38	25	15.2	26	2 A34280	phospholipase A2 (
39	25	15.2	27	2 B44180	phospholipase A2 (
40	25	15.2	30	2 PC2361	alpha-glucosidase
41	25	15.2	30	2 S21815	H+-transporting AT
42	25	15.2	30	2 D72276	hypothetical prote
43	24	14.6	17	2 S15778	insulin chain B -
44	24	14.6	17	2 A58946	formylmethanofuran
45	24	14.6	18	2 S71592	serine proteinase

ALIGNMENTS

RESULT 1  
I77372  
CD44SP - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I77372  
R: Tanabe, K.K.; Nishi, T.; Sava, H.  
Mol. Carcinog. 7, 212-220, 1993  
A:Title: Novel variants of CD44 arising from alternative splicing: changes in the CD4  
A:Reference number: I57483; MUID:93356912  
A:Accession: I77372  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-29 <RES>  
A:Cross-references: GB:S66400; NID:g435697; PIDN:ARB27917.1; PID:g435698  
C:Genetics:  
A:Gene: GDB:CD44  
A:Cross-references: GDB:I20739; OMIM:107269  
A:Map position: 11pter-11p13  
A:Introns: 23/1

Query Match 32.3%; Score 53; DB 2; Length 29;  
Best Local Similarity 100.0%; Pred.No.0.26;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LCLVPLSLAQI 11  
|||||  
Db 12 LCLVPLSLAQI 22

RESULT 2  
S78310  
hypothetical protein 29a - Odontella sinensis chloroplast  
C:Species: Chloroplast Odontella sinensis  
C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 24-Apr-1998  
C:Accession: S78310  
R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.  
Plant Mol. Biol. Rep. 13, 336-342, 1995  
A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sine  
A:Reference number: S78238  
A:Accession: S78310  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-29 <KOW>  
A:Cross-references: EMBL:Z67753; NID:g1185127; PID:e211861; PID:g1185200  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 20.7%; Score 34; DB 2; Length 29;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LCLVPLSLAQIDLN 14  
|||::|||  
Db 12 LCLFCTNIDQIDFN 25

## RESULT 3

S00579  
hypothetical protein 547 (hisa 5' region) - Methanococcus vannielii (fragment)  
C:Species: Methanococcus vannielii  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 22-Oct-1999  
C:Accession: S00579  
R:Brown, J.W.; Thomm, M.; Beckler, G.S.; Frey, G.; Stetter, K.O.; Reeve, J.N.  
Nucleic Acids Res. 16, 135-150, 1988  
A:Title: An archaeobacterial RNA polymerase binding site and transcription initiation of  
A:Reference number: S00579; MUID:88124191  
A:Accession: S00579  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <BR>  
A:Cross-references: EMBL:X07391; MID:g44714; PIDN:CAA30299.1; PID:g44715

Query Match 19.5%; Score 32; DB 2; Length 26;  
Best Local Similarity 85.7%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 VERNGRY 31  
||| |||  
Db 19 VERGRY 25

## RESULT 4

F44957  
osmotin homolog - common tobacco (cv. Samsun NN) (fragment)  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 12-Apr-1995  
C:Accession: F44957  
R:Takeda, S.; Sato, F.; Ida, K.; Yamada, Y.  
Plant Cell Physiol. 31, 215-221, 1990  
A:Title: Characterization of polypeptides that accumulate in cultured Nicotiana tabacum  
A:Reference number: A44957  
A:Accession: F44957  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <TAK>  
C:Superfamily: thaumatin I

Query Match 18.9%; Score 31; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 4e+02;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 20 AGVFHVKNRGY 31  
:||| | |  
Db 1 SGVFVHNXPY 12

## RESULT 5

A60889  
olfactory glycoprotein RB-8 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: A60889  
R:Schwab, J.E.; Gottlieb, D.I.  
J. Neurosci. 8, 3470-3480, 1988  
A:Title: Purification and characterization of an antigen that is spatially segregated in  
A:Reference number: A60889; MUID:89010968  
A:Accession: A60889  
A:Molecule type: protein

A:Residues: 1-17 <SCH>  
A:Note: This protein was purified from whole brain  
C:Comment: The monoclonal antibody RB-8 binds this integral membrane glycoprotein on  
C:Keywords: glycoprotein; membrane protein; olfaction

Query Match 18.3%; Score 30; DB 2; Length 17;  
Best Local Similarity 33.3%; Pred. No. 4.4e+02;  
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 LCLVPLSLAQIDLNI 15  
| | |::| | |  
Db 1 LLXVTISLSKVELSV 15

## RESULT 6

S10052  
ribosomal protein L37.e - fission yeast (Schizosaccharomyces pombe) (fragment)  
N:Alternate names: ribosomal protein SP-L27  
C:Species: Schizosaccharomyces pombe  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 30-Sep-1993  
C:Accession: S10052  
R:Otaka, E.; Higo, K.I.; Itoh, T.  
Mol. Gen. Genet. 191, 519-524, 1983  
A:Title: Yeast ribosomal proteins: VII. Cytoplasmic ribosomal proteins from Schizosaccharomyces pombe  
A:Reference number: S07293; MUID:84038947  
A:Accession: S10052  
A:Molecule type: protein  
A:Residues: 1-28 <OTA>  
C:Superfamily: rat ribosomal protein L37  
C:Keywords: protein biosynthesis; ribosome

Query Match 18.3%; Score 30; DB 2; Length 28;  
Best Local Similarity 60.0%; Pred. No. 7.2e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 18 RFAGVFHVBEK 27  
| | | | |  
Db 19 RRGKGFHVOK 28

## RESULT 7

A43826  
dnaK-type molecular chaperone - Ajellomyces capsulata (fragment)  
N:Alternate names: 80K heat shock protein 70 homolog  
C:Species: Ajellomyces capsulata, Histoplasma capsulatum  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 17-Mar-1999  
C:Accession: A43826  
R:Gomez, F.J.; Gomez, A.M.; Deepe Jr., G.S.  
Infect. Immun. 60, 2565-2571, 1992  
A:Title: An 80-kilodalton antigen from Histoplasma capsulatum that has homology to he  
A:Reference number: A43826; MUID:92307875  
A:Accession: A43826  
A:Molecule type: protein  
A:Residues: 1-18 <GOM>  
A:Experimental source: extract of cell wall and cell membrane  
C:Function:  
A:Description: involved in protein folding and assembling/disassembling of protein co  
C:Superfamily: heat shock protein 70  
C:Keywords: ATP; heat shock; molecular chaperone; stress-induced protein

Query Match 17.7%; Score 29; DB 2; Length 18;  
Best Local Similarity 46.2%; Pred. No. 6.5e+02;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 11 IDLNITCFAGVF 23  
||| | |  
Db 6 IDLGTYSVVGIF 18

## RESULT 8

H64711  
hypothetical protein HP1536 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: H64711  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: H64711  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-18 <TOM>  
A:Cross-references: GB:AE000651; GB:AE000511; NID:g2314708; PIDN:AAD08578.1; PID:g2314708

Query Match 17.7%; Score 29; DB 2; Length 18;  
Best Local Similarity 40.0%; Pred. No. 6.5e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 17 CRFAGVHFHVEKNRGY 31  
||| | : | : |  
Db 5 CRFLN--HPSRNPKY 17

RESULT 9  
I45911  
dnaK-type molecular chaperone hsp70A - bovine (fragment)  
N:Alternate names: heat shock protein  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 20-Aug-1999  
C:Accession: I45911  
R:Kowalski, J.; Gilbert, S.A.; van Drunen-Littel-van den Hurk, S.; van den Hurk, J.; Bab Vaccine 11, 1100-1107, 1993  
A:Title: Heat-Shock Promoter-Driven Synthesis of Secreted Herpesvirus Glycoproteins in T  
A:Reference number: I45911; MUID:94070117  
A:Accession: I45911  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-28 <KOW>  
A:Cross-references: GB:M98823; NID:g409185; PIDN:AAA30568.1; PID:g409186  
C:Genetics:  
A:Gene: hsp70A  
C:Function:  
A:Description: involved in protein folding and assembling/disassembling of protein comp  
C:Superfamily: heat shock protein 70  
C:Keywords: ATP; molecular chaperone

Query Match 17.7%; Score 29; DB 2; Length 28;  
Best Local Similarity 53.8%; Pred. No. 1e+03;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 11 IDLNITCRFAGVF 23  
||| | : | : |  
Db 9 IDLGTYSVGVGF 21

RESULT 10  
B82233  
hypothetical protein VC1175 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: B82233  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Emlaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H i, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833

A:Accession: B82233  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-30 <HEI>  
A:Cross-references: GB:AE004197; GB:AE003852; NID:g9655642; PIDN:AAF94334.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1175  
A:Map position: 1

Query Match 17.7%; Score 29; DB 2; Length 30;  
Best Local Similarity 46.2%; Pred. No. 1.1e+03;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 19 FAGVHFHVEKNRGY 31  
| | | : | : |  
Db 5 FIGFPHLSLLFRY 17

RESULT 11  
S31613  
beta-1,3-glucanase homolog (clone A28) - rape (fragment)  
C:Species: Brassica napus (rape)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 17-Nov-2000  
C:Accession: S31613  
R:Hird, D.; Worrall, D.; Hodge, R.; Paul, W.; Smartt, S.; Draper, J.; Scott, R. submitted to the EMBL Data Library, December 1992  
A:Description: The auther-specific protein encoded by the Brassica napus and Arabidop  
A:Reference number: S31612  
A:Accession: S31613  
A:Molecule type: mRNA  
A:Residues: 1-19 <HIR>  
A:Cross-references: EMBL:X69890; NID:g17735; PID:g17736  
A:Experimental source: clone A28  
C:Superfamily: beta-1,3-glucanase

Query Match 17.1%; Score 28; DB 2; Length 19;  
Best Local Similarity 55.6%; Pred. No. 9.7e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 14 NITCRFAGV 22  
| | | : | : |  
Db 9 NDRCKFPQGV 17

RESULT 12  
A83698  
hypothetical protein BH0385 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: A83698  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20263314  
A:Accession: A83698  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-23 <STO>  
A:Cross-references: GB:AF001508; GB:BA000004; NID:g10172890; PIDN:BA804104.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0385

Query Match 17.1%; Score 28; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GVFW 25  
|||||

Db 3 GVPHV 7

## RESULT 13

A81078

hypothetical protein NMB1491 [imported] - Neisseria meningitidis (strain MC58 serogroup C)  
 C:Species: Neisseria meningitidis  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: A81078  
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ric, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000  
 A:Authors: Grandl, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, R.; Tettelin, H.; Vamathevan, J.; Venter, J.C.; Fraser, C.M.  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755  
 A:Accession: A81078  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-29 <TET>  
 A:Cross-references: GB:AE002498; GB:AE002098; NID:g7226724; PIDN:AAF41847.1; PID:g7226724  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1491

Query Match 17.1%; Score 28; DB 2; Length 29;

Best Local Similarity 71.4%; Pred. No. 1.5e+03;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LCLVPLS 7

| | | : | |

Db 9 LCLMPAS 15

## RESULT 14

D82348

hypothetical protein VC0226 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C:Species: Vibrio cholerae  
 C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: D82348  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Church, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R. L., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833  
 A:Accession: D82348  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-31 <HET>  
 A:Cross-references: GB:AE004112; GB:AE003852; NID:g9654625; PIDN:AAF93402.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC0226  
 A:Map position: 1

Query Match

Best Local Similarity 35.7%; Score 28; DB 2; Length 31;

Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 17 CRFAGVPHVEKNGR 30

| : | | |

Db 4 CSASGVLHFTFSQG 17

## RESULT 15

B61525

beta-lactoglobulin - crab-eating macaque (fragment)  
 C:Species: Macaca fascicularis (crab-eating macaque)  
 C>Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 11-May-2000  
 C:Accession: B61525

R:Azuma, N.; Yamauchi, K.  
 Comp. Biochem. Physiol. B 99, 917-921, 1991  
 A:Title: Identification of alpha-lactalbumin and beta-lactoglobulin in cynomolgus monkey  
 A:Reference number: A61525; MUID:92164307  
 A:Accession: B61525

A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-26 <AZU>  
 C:Superfamily: lipocalin; lipocalin homology

Query Match 16.8%; Score 27.5; DB 2; Length 26;

Best Local Similarity 30.0%; Pred. No. 1.6e+03;

Matches 6; Conservative 5; Mismatches 6; Indels 3; Gaps 1;

QY 5 PLSLAQIDILNITCRFAGVPH 24

| : : | | | : | | : |

Db 4 POTMODVDLP---KLAGTWH 20

Search completed: February 21, 2002, 17:09:01  
 Job time: 112 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:08:44 ; Search time 10.06 Seconds  
(without alignments)  
112.983 Million cell updates/sec

Title: US-08-753-851-12  
Perfect score: 164  
Sequence: 1 LCLVPLSLAQIDLNITCFAGVFHVEKNGRY 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1886

Minimum DB seq length: 0  
Maximum DB seq length: 31

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	20.7	29	1 YCX4 ODOSI	P49830 odontella s
2	27	16.5	30	1 DEF2 MACMU	P82317 macaca mula
3	26.5	16.2	31	1 PETL CHLVU	P56306 chorella v
4	26	15.9	15	1 UC29 MAIZE	P80635 zea mays (m
5	26	15.9	20	1 EMA2 MOUSE	P31230 mus musculu
6	26	15.9	21	1 CFPA TREPH	P56738 treponema p
7	25.5	15.5	21	1 YD90 HAEIN	P45194 haemophilus
8	25	15.2	30	1 VAA2 EQUAR	Q04238 equisetum a
9	24	14.6	19	1 FIBE HORSE	P14471 equus caball
10	24	14.6	25	1 PETM SRECO	O96810 skeletonema
11	24	14.6	30	1 SCK2 TITSE	P08816 titiys serr
12	24	14.6	31	1 CXMA CONMR	P56708 conus marmo
13	24	14.6	31	1 Y3KD BPCHP	P19187 bacterioph
14	23	14.0	16	1 LE05 BIOGL	P80744 biomphalari
15	23	14.0	19	1 LCRP PETMA	Q10996 petromyzon
16	23	14.0	19	1 LPGE ECOLI	P33236 escherichia
17	23	14.0	26	1 MGN CHICK	P50594 gallus gall
18	23	14.0	31	1 DIUX DIPPU	P82372 diptoptera
19	23	14.0	31	1 PETN CYACA	Q91lr6 cyanidium c
20	22.5	13.7	28	1 PA2C PSEPO	P20260 pseudochis
21	22.5	13.7	29	1 BR2D RANES	P40840 rana escul
22	22	13.4	8	1 UF06 MOUSE	P38644 mus musculu
23	22	13.4	14	1 NEJ2 FASHE	P80526 fasciola he
24	22	13.4	24	1 SCX2 MESTA	P45668 mesobuthus
25	22	13.4	28	1 HOX1 ARCFU	O30163 archaeoglob
26	22	13.4	29	1 HOX1 RHOOP	P22660 rhodococcus
27	22	13.4	30	1 PSAM CYACA	O91lx5 cyanidium c
28	22	13.4	31	1 DEF2 MESAU	P81466 mesocricetu
29	21.5	13.1	24	1 RS13 THETH	P80377 thermus aqu
30	21.5	13.1	24	1 THD2 SALTY	P11954 salmonella
31	21.5	13.1	28	1 PHYB ASPFI	P81440 aspergillus
32	21.5	13.1	29	1 BREE RANES	P40841 rana escul
33	21.5	13.1	31	1 PSAM CHLVU	P56314 chorella v

34 21 12.8 9 1 FIBB PAPH

35 21 12.8 12 1 UKA2 HUMAN

36 21 12.8 14 1 XA34 MAIZE

37 21 12.8 15 1 NIS1 ANASQ

38 21 12.8 17 1 FLAW AZOCH

39 21 12.8 18 1 YPE4 LACLC

40 21 12.8 19 1 COOT SARBU

41 21 12.8 19 1 PTHP STRSA

42 21 12.8 20 1 FLAW AZOVI

43 21 12.8 22 1 DNIV BPDIO

44 21 12.8 25 1 FLB1 TREHY

45 21 12.8 25 1 NEUU PIG

P19343 papio hamad

P31144 homo sapien

P80640 zea mays (m

Q44507 anabaena sp

P23002 azotobacter

P42022 lactococcus

Q09148 sarcophaga

P24365 streptococ

P52964 azotobacter

Q38199 bacterioph

P80158 treponema h

P34964 sus scrofa

#### ALIGNMENTS

RESULT 1

YCX4\_ODOSI STANDARD; PRT; 29 AA.

AC P49830;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE HYPOTHETICAL 3.5 KDA PROTEIN IN YCF33-TRNY INTERGENIC REGION (ORF29A).

OS Odontella sinensis.

OG Chloroplast.

OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;

OC Biddulphiophycidae; Eupodiscales; Eupodiscaeae; Odontella.

OX NCBI\_TaxID=2839;

RN [1]

RP SEQUENCE FROM N.A.

RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;

RT "The chloroplast genome of a chlorophyll a+c-containing alga,

RT Odontella sinensis.";

RL Plant Mol. Biol. Rep. 13:336-342(1995).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; Z67753; CAA91683.1; -

KW Chloroplast: Hypothetical protein.

SQ SEQUENCE 29 AA; 3481 MW; 4960F2ACA72ACBA2 CRC64;

Query Match 20.7%; Score 34; DB 1; Length 29;

Best Local Similarity 50.0%; Pred. No. 67;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LCLVPLSLAQIDLN 14

||| : |||

Db 12 LCLFCTNIDQIDFN 25

RESULT 2

DEF2\_MACMU STANDARD; PRT; 30 AA.

AC P82317;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE NEUTROPHIL DEFENSIN 2 (RMAD-2).

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Macaca.

OX NCBI\_TaxID=9544;

RN [1]





```

RP MEDLINE=93015897; PubMed=1400342;
RA Kao J., Ryan J., Brett G., Chen J., Shen H., Pan Y.-G., Godman G.,
RA Familletti P.C., Wang F., Pan Y.-C.E., Stern D., Claus M.;
RT "Endothelial monocyte-activating polypeptide II. A novel
RT tumor-derived polypeptide that activates host-response mechanisms.";
RL J. Biol. Chem. 267:20239-20247(1992).
CC -!- FUNCTION: ALTERS ENDOTHELIAL AND MONOCYTE FUNCTIONS. INDUCES THE
CC MIGRATION OF MONOCYTES AND GRANULOCYTES, AND INDUCES AN
CC INFLAMMATORY RESPONSE IN THE MOUSE FOOTPAD MODEL. EMAP II ELICITS
CC A PHLOGOGENIC RESPONSE AND, POTENTIALLY, AUGMENTS THE EFFECTS OF
CC THE OTHER TUMOR-DERIVED CYTOKINES.
CC -!- SUBUNIT: MONOMER.
CC PIR: A44032; A44032.
DR Cytokine.
KW VARIANT
FT NON_TER
SQ SEQUENCE 20 AA; 2164 MW; 018E08BBED066721 CRC64;

Query Match 15.9%; Score 26; DB 1; Length 20;
Best Local Similarity 36.4%; Pred. No. 7.6e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 PLSLAQIDLNI 15
Db 3 PIDASRLDLRI 13

RESULT 6
ID CFPA_TREPH STANDARD; PRT; 21 AA.
AC P56738;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOPLASMIC FILAMENT PROTEIN A (FRAGMENT).
GN CFPA.
OS Treponema phagedenis.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=162;
RN [1]
RP SEQUENCE.
RC STRAIN=KAZAN 5;
RX MEDLINE=96236033; PubMed=8655496;
RA You Y., Elmore S., Colton L.L., Mackenzie C., Stoops J.K.,
RA Weinstock G.M., Norris S.J.;
RT "Characterization of the cytoplasmic filament protein gene (cfpa) of
RT Treponema pallidum subsp. pallidum.";
RL J. Bacteriol. 178:3177-3187(1996).
CC -!- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
CC LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
CC -!- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
CC APPosition TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
CC UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
KW Structural protein; Antigen.
FT NON_TER
SQ SEQUENCE 21 AA; 2231 MW; 574604B4FFC2D017 CRC64;

Query Match 15.9%; Score 26; DB 1; Length 21;
Best Local Similarity 83.3%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 VFHVEK 27
Db 11 VFHPEK 16

RESULT 7
ID YD90_HABIN STANDARD; PRT; 21 AA.
AC P45194;

```

```

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN HI1390.
GN HI1390.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Banton R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32819; AAC23037.1;
DR TIGR: HI1390;
DR InterPro: IPR001109; HupF_HypC.
DR PROSITE: PS01097; HUPF_HYPC; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 21 AA; 2387 MW; 2A63C08B47C2ED3B CRC64;

Query Match 15.5%; Score 25.5; DB 1; Length 21;
Best Local Similarity 41.7%; Pred. No. 9.5e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

QY 1 LCL-VPLSLAQI 11
Db 1 MCLGVPIKLSKL 12

RESULT 8
ID VAA2_EQUAR STANDARD; PRT; 30 AA.
AC Q04238;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A, ISOFORM 2 (EC 3.6.1.34)
DE (FRAGMENT).
OS Equisetum arvense (Field horsetail) (Common horsetail).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Equisetophyta; Sphenopsida; Equisetales; Equisetaceae; Equisetum.
OX NCBI_TaxID=3258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93138084; PubMed=8422915;
RA Starke T., Gogarten J.P.;
RT "A conserved intron in the V-ATPase A subunit genes of plants and
RT algae.";
RL FEBS Lett. 315:252-258(1993).
CC -!- FUNCTION: CATALYTIC SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF
CC VACUOLAR ATPASE. V-ATPASE VACUOLAR ATPASE IS RESPONSIBLE FOR

```

```

KW Blood coagulation; Plasma; Sulfation.
FT MOD_RES 3 3 SULFATION.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2296 MW; 921A2B02D5F6891D CRC64;

Query Match 14.6%; Score 24; DB 1; Length 19;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 24 HVEKNR 30
DB 5 HEEDGR 11

RESULT 10
PETM_SKECO STANDARD; PRT; 25 AA.
ID PETM_SKECO
AC 096810;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME B6-F COMPLEX SUBUNIT VII (CYTOCHROME B6F COMPLEX SUBUNIT
DE PETM) (FRAGMENT).
DE GN PETM OR YCF31.
OS Skeletonema costatum.
OC Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Thalassiosirophyceidae; Thalassiosirales; Skeletonemataceae;
OC Skeletonema.
OX NCBI_TaxID=2843;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=NIES-323;
RA Tada N., Otsuka S., Oyaizu H., Matsumoto S.;
RL "Plastid DNA sequences of Skeletonema costatum NIES 323."
RT Submitted (JAN-1999) to the EMBL/GenBank/DBDJ databases.
CC -1- FUNCTION: CYTOCHROME B6-F COMPLEX SUBUNIT.
CC -1- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE-ASSOCIATED (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PETM FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; AJ132266; CA10631.1; -.
Electron transport; Chloroplast; Respiratory chain; Thylakoid;
Transmembrane.
FT NON_TER 1 1
FT TRANSMEM 2 22 POTENTIAL.
SQ SEQUENCE 25 AA; 2630 MW; 99ECF1B6CA3113AA CRC64;

Query Match 14.6%; Score 24; DB 1; Length 25;
Best Local Similarity 44.4%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 ITCRFAGVF 23
DB 3 VTCIFMTLF 11

RESULT 11
SCK2_TITSE STANDARD; PRT; 30 AA.
ID SCK2_TITSE
AC P08816;
DT 01-NOV-1988 (Rel. 09, Created sequence update)
DT 01-NOV-1988 (Rel. 09, Last sequence update)

```

DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE TOXIN II-9 (FRAGMENT).  
 OS Tityus serrulatus (Brazilian scorpion).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 OC Buthoidea; Buthidae; Tityus.  
 OX NCBI\_TaxID=6887;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Venom;  
 RA Possani L.D., Martin B.M., Svendsen I.;  
 RT "The primary structure of noxiustoxin, A K channel blocking peptide,  
 purified from the venom of the scorpion Centruroides noxius  
 Hoffmann.";  
 RL Carlsberg Res. Commun. 47:285-289(1982).  
 CC -!- FUNCTION: BLOCKER FOR CALCIUM-ACTIVATED POTASSIUM CHANNELS.  
 CC -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY. POTASSIUM  
 CHANNEL INHIBITORS SUBFAMILY.  
 DR HSP: P24662; LKTX.  
 DR InterPro: IPR001947; Charybdotoxin.  
 DR Pfam: PF00451; toxin.2; 1.  
 DR PROSITE: PS01138; SCORP\_SHORT\_TOXIN; PARTIAL.  
 KW Neurotoxin; Potassium channel inhibitor.  
 FT DISULFID 7 28 BY SIMILARITY.  
 FT DISULFID 13 ? BY SIMILARITY.  
 FT DISULFID 17 ? BY SIMILARITY.  
 FT VARIANT 21 21 I -> F.  
 FT VARIANT 29 29 V -> M.  
 FT NON\_TER 30 30  
 FT SEQUENCE 30 AA; 3164 MW; 19196GEDBB136DFE CRC64;

Query Match 14.6%; Score 24; DB 1; Length 30;

Best Local Similarity 50.0%; Pred. No. 2.3e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 13 LNITCR 18  
 :|: |  
 DB 3 INVKCR 8

RESULT 12  
 CXMA\_CONMR STANDARD; PRT; 31 AA.  
 ID P56708;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE MU-O-CONOTOXIN MRVIA.  
 OS Conus marmoreus (Marble cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=42752;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Venom;  
 RX MEDLINE=95348106; PubMed=7622492;  
 RA McIntosh J.M., Hasson A., Spira M.E., Gray W.R., Li W., Marsh M.,  
 RA Hillyard D.R., Olivera B.M.;  
 RT "A new family of conotoxins that blocks voltage-gated sodium  
 channels.";  
 RL J. Biol. Chem. 270:16796-16802(1995).  
 CC -!- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
 MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
 CHANNELS.  
 CC -!- MASS SPECTROMETRY: MW=3487.8; METHOD=LSIMS.  
 KW Sodium channel inhibitor; Venom.  
 FT DISULFID 2 20 BY SIMILARITY.  
 FT DISULFID 9 25 BY SIMILARITY.  
 FT DISULFID 19 30 BY SIMILARITY.  
 FT SEQUENCE 31 AA; 3495 MW; 741FA610E6F9D289 CRC64;

Query Match 14.6%; Score 24; DB 1; Length 31;

Best Local Similarity 60.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CLVPL 6  
 :|: |  
 DB 9 CIVPI 13

RESULT 13  
 Y3KD\_BPCHP  
 ID Y3KD\_BPCHP STANDARD; PRT; 31 AA.  
 AC P19187;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-NOV-1990 (Rel. 16, Last annotation update)  
 DE HYPOTHETICAL 3.6 KDA PROTEIN (ORF7).  
 OS Bacteriophage Chp1.  
 OS Viruses.  
 OX NCBI\_TaxID=12367;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90111716; PubMed=2607341;  
 RA Storey C.C., Lusher M., Richmond S.J.;  
 RT "Analysis of the complete nucleotide sequence of Chp1, a phage which  
 infects avian Chlamydia psittaci.";  
 RL J. Gen. Virol. 70:3381-3390(1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL; D00624; BAA00504.1; -  
 DR PIR; J00351; J00351.  
 KW Hypothetical protein.  
 SQ SEQUENCE 31 AA; 3575 MW; 2F52D99439F181D3 CRC64;

Query Match 14.6%; Score 24; DB 1; Length 31;

Best Local Similarity 22.2%; Pred. No. 2.4e+03;  
 Matches 4; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 LCLVPLSLAQIDLNITCR 18  
 :|: :|: :|: |  
 DB 1 MCLLVKSVMLSVISLR 18

RESULT 14  
 LE05\_BIOGL STANDARD; PRT; 16 AA.  
 ID LE05\_BIOGL  
 AC P80744;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HEMOLYMPH 65 KDA LECTIN BG05 (FRAGMENT).  
 GN BG05.  
 OS Biomphalaria glabrata (Bloodfluke planorb).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Basommatophora;  
 OC Planorbidae; Biomphalaria.  
 OX NCBI\_TaxID=6526;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=M-LINE; TISSUE=Hemolymph;  
 RX MEDLINE=97385165; PubMed=9238039;  
 RA Adema C.M., Hertel L.A., Miller R.D., Loker E.S.;  
 RT "A family of fibrinogen-related proteins that precipitates parasite-  
 RT derived molecules is produced by an invertebrate after infection.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:8691-8696(1997).  
 CC -!- FUNCTION: BINDS AND PRECIPITATES ANTIGENS OF THE PARASITE  
 CC ECHINOSTOMA PARAENSEI.

CC -!- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.  
CC -!- INDUCTION: BY INFECTION.

KW Lectin.  
FT NON\_TER 1 1  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1790 MW; 57489A8F2EEDAA94 CRC64;

Query Match 14.0%; Score 23; DB 1; Length 16;  
Best Local Similarity 45.5%; Pred. No. 1.7e+03;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 LAQIDLNITCR 18  
||| :|||  
Db 6 LAQYVDLTAR 16

## RESULT 15

LCRP\_PETMA STANDARD; PRT; 19 AA.  
AC Q10996;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE CORTICOSTATIN-RELATED PROTEIN LCRP.  
OS Petromyzon marinus (Sea lamprey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
OX NCBI\_TaxID=7757;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Skin;  
RX MEDLINE=96321324; PubMed=8759287;  
RA Conlon J.M., Sower S.A.;  
RT "Isolation of a peptide structurally related to mammalian  
corticostatins from the lamprey Petromyzon marinus.";  
RL Comp. Biochem. Physiol. 114B:133-137(1996).  
CC -!- FUNCTION: MAY HAVE MICROBICIDAL ACTIVITIES. MAY INHIBIT  
CORTICOTROPIN (ACTH) STIMULATED STEROIDOGENESIS AND THE MICROBIAL  
ACTIONS OF THE CORTICOSTATINS.  
CC -!- MASS SPECTROMETRY: MW=2201; MW\_ERR=0.4; METHOD=ELECTROSPRAY.  
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF CORTICOSTATIN/DEFENSIN  
PEPTIDES.  
KW Defensin; Antibiotic.  
FT DISULFID 1 18 BY SIMILARITY.  
FT DISULFID 3 9 BY SIMILARITY.  
FT DISULFID 8 17 BY SIMILARITY.  
SQ SEQUENCE 19 AA; 2209 MW; 8D9CDC71A199AE5 CRC64;

Query Match 14.0%; Score 23; DB 1; Length 19;  
Best Local Similarity 57.1%; Pred. No. 2e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 LNVTCRF 19  
||: |||  
Db 13 LNVYCCF 19

Search completed: February 21, 2002, 17:11:46  
Job time: 182 sec

Result No.	Query	Score	Match	Length	DB	ID	Description
1	37	22.6	28	1	Q9UXR3	Q9uxr3 methanosarc	
2	34	20.7	27	12	Q9QH74	Q9qh74 hepatitis c	
3	32	19.5	20	5	Q9TWG8	Q9twg8 aedes aegypt	
4	31	19.5	26	1	Q50835	Q50835 methanococc	
5	31	18.9	26	10	Q9S8K5	Q9s8k5 nicotiana t	
6	31	18.9	27	12	Q9QH83	Q9qh83 hepatitis c	
7	31	18.9	27	12	Q9QH82	Q9qh82 hepatitis c	
8	31	18.9	27	12	Q9QH81	Q9qh81 hepatitis c	
9	31	18.9	27	12	Q9QH80	Q9qh80 hepatitis c	
10	31	18.9	27	12	Q9QH79	Q9qh79 hepatitis c	
11	31	18.9	27	12	Q9QH78	Q9qh78 hepatitis c	
12	31	18.9	27	12	Q9QH77	Q9qh77 hepatitis c	
13	31	18.9	27	12	Q9QH76	Q9qh76 hepatitis c	
14	31	18.9	27	12	Q9QH75	Q9qh75 hepatitis c	
15	31	18.9	27	12	Q9QH73	Q9qh73 hepatitis c	
16	31	18.9	27	12	Q9QH72	Q9qh72 hepatitis c	
17	31	18.9	27	12	Q9QH70	Q9qh70 hepatitis c	
18	31	18.9	27	12	Q9QH69	Q9qh69 hepatitis c	
19	31	18.9	27	12	Q9QH68	Q9qh68 hepatitis c	

QY	13	LNITCRFAGVFHV	25
	:	: : : :	:
DB	7	VNETAQFSSIAHI	19

RESULT 6  
Q9QH83  
ID Q9QH83 PRELIMINARY; PRT: 27 AA.

AC Q9QH83:  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPEPTIDE [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
OS (FRAGMENT).  
DE Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
RT genome and the sensitivity of hepatitis C virus to interferon alpha  
RT therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
CC CONTAINED WITHIN THE GENOME POLYPEPTIDE THAT CONTAINS: CAPSID  
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
CC FAMILY.  
DR EMBL; AF167045; AAD53678.1; -.  
DR InterPro; IPR002531; HCV\_NSI.  
DR Pfam; PF01560; HCV\_NSI.1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 7 SLAQIDLNTICRFAGVHFVEKNGR 30  
DB 4 TVGSISSRATASFAGLSLGSGQR 27

RESULT 7  
Q9QH82  
ID Q9QH82 PRELIMINARY; PRT; 27 AA.  
AC Q9QH82;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPEPTIDE [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
RT genome and the sensitivity of hepatitis C virus to interferon alpha  
RT therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
CC CONTAINED WITHIN THE GENOME POLYPEPTIDE THAT CONTAINS: CAPSID  
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1

CC FAMILY.  
DR EMBL; AF167046; AAD53679.1; -.  
DR InterPro; IPR002531; HCV\_NSI.  
DR Pfam; PF01560; HCV\_NSI.1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 7 SLAQIDLNTICRFAGVHFVEKNGR 30  
DB 4 TVGSISSRATASFAGLSLGSGQR 27

RESULT 8  
Q9QH81  
ID Q9QH81 PRELIMINARY; PRT; 27 AA.  
AC Q9QH81;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPEPTIDE [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
RT genome and the sensitivity of hepatitis C virus to interferon alpha  
RT therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
CC CONTAINED WITHIN THE GENOME POLYPEPTIDE THAT CONTAINS: CAPSID  
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
CC FAMILY.  
DR EMBL; AF167047; AAD53680.1; -.  
DR InterPro; IPR002531; HCV\_NSI.  
DR Pfam; PF01560; HCV\_NSI.1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 7 SLAQIDLNTICRFAGVHFVEKNGR 30  
DB 4 TVGSISSRATASFAGLSLGSGQR 27

RESULT 9  
Q9QH80  
ID Q9QH80 PRELIMINARY; PRT; 27 AA.  
AC Q9QH80;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
 DE (FRAGMENT).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
 RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
 RT genome and the sensitivity of hepatitis C virus to interferon alpha  
 RT therapy.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
 CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
 CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
 CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
 CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
 CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
 CC FAMILY.  
 DR EMBL; AF167048; AAD53681.1; -.  
 DR InterPro; IPR002531; HCV\_NSL.  
 DR Pfam; PF01560; HCV\_NSL; 1.  
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
 Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
 Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 7 SLAQIDLNITCRFAGVHVEKNR 30  
 : : | | | : |  
 Db 4 TVGSISSRATASFAGLSGSGQR 27

RESULT 10  
 OYQH79 PRELIMINARY; PRT; 27 AA.  
 AC OYQH79;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
 DE (FRAGMENT).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
 RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
 RT genome and the sensitivity of hepatitis C virus to interferon alpha  
 RT therapy.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
 CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
 CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
 CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
 CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
 CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
 CC FAMILY.  
 DR EMBL; AF167049; AAD53682.1; -.

DR InterPro; IPR002531; HCV\_NSL.  
 DR Pfam; PF01560; HCV\_NSL; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
 Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
 Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 7 SLAQIDLNITCRFAGVHVEKNR 30  
 : : | | | : |  
 Db 4 TVGSISSRATASFAGLSGSGQR 27

RESULT 11  
 OYQH78 PRELIMINARY; PRT; 27 AA.  
 AC OYQH78;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
 DE (FRAGMENT).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
 RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
 RT genome and the sensitivity of hepatitis C virus to interferon alpha  
 RT therapy.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
 CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
 CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
 CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
 CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
 CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
 CC FAMILY.  
 DR EMBL; AF167050; AAD53683.1; -.  
 DR InterPro; IPR002531; HCV\_NSL.  
 DR Pfam; PF01560; HCV\_NSL; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
 Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
 Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 7 SLAQIDLNITCRFAGVHVEKNR 30  
 : : | | | : |  
 Db 4 TVGSISSRATASFAGLSGSGQR 27

RESULT 12  
 OYQH77 PRELIMINARY; PRT; 27 AA.  
 AC OYQH77;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)



DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Andres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
RT genome and the sensitivity of hepatitis C virus to interferon alpha  
RT therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
CC FAMILY.  
CC EMBL; AF167051; AAD53684.1; -.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01560; HCV\_NSI; 1.  
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 7 SLAQIDLNITCRFAGVHFVEKNR 30  
: : | | | | :  
Db 4 TVGSISSRATASFAGLSLGSOOR 27

RESULT 13  
ID Q9QH76 PRELIMINARY; PRT; 27 AA.  
AC Q9QH76;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Andres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
RT genome and the sensitivity of hepatitis C virus to interferon alpha  
RT therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
CC FAMILY.  
CC EMBL; AF167052; AAD53685.1; -.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01560; HCV\_NSI; 1.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 7 SLAQIDLNITCRFAGVHFVEKNR 30  
: : | | | | :  
Db 4 TVGSISSRATASFAGLSLGSOOR 27

RESULT 14  
ID Q9QH75 PRELIMINARY; PRT; 27 AA.  
AC Q9QH75;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Andres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of the hypervariable region 1 of the viral  
RT genome and the sensitivity of hepatitis C virus to interferon alpha  
RT therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NSF1 IS  
CC CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
CC PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
CC E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
CC NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPOPROTEIN ENVELOPE  
CC THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
CC FAMILY.  
CC EMBL; AF167053; AAD53686.1; -.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01560; HCV\_NSI; 1.  
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2754 MW; D4385958667D5361 CRC64;

Query Match 18.9%; Score 31; DB 12; Length 27;  
Best Local Similarity 29.2%; Pred. No. 7.8e+02;  
Matches 7; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 7 SLAQIDLNITCRFAGVHFVEKNR 30  
: : | | | | :  
Db 4 TVGSISSRATASFAGLSLGSOOR 27

RESULT 15  
ID Q9QH73 PRELIMINARY; PRT; 27 AA.  
AC Q9QH73;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:05:14 ; Search time 23.51 Seconds  
(without alignments)  
97.672 Million cell updates/sec

Title: US-08-753-851-12

Perfect score: 164

Sequence: 1 LCLVPLSLAQIDENITCRFAGVFHVKNGRY 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 240333

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	164	100.0	31	15 AAR53482	CD44 peptide CD44-
2	69	42.1	25	15 AAR53496	Human CD44H N-term
3	58	35.4	25	15 AAR53495	African green monk
4	43	26.2	30	9 AAP82919	Product from plasm
5	42.5	25.9	31	9 AAP82917	Product from plasm
6	38	23.2	27	19 AAW47362	Clone CC397.19, H
7	38	23.2	27	19 AAW42020	Clone CC397-11 pro
8	38	23.2	30	20 AAY08628	Human secreted pro
9	38	23.2	30	21 AAB34787	Human secreted pro
10	38	23.2	30	21 AAY67319	Human secreted pro
11	37	22.6	23	22 AAE02383	Human alpha1-antit

12	37	22.6	29	9 AAP82918	Product from plasm
13	36	22.0	21	19 AAW67044	Laminin peptide an
14	36	22.0	26	22 AAB84641	Antigenic peptide
15	36	22.0	28	20 AAY19691	SEQ ID NO 409 from
16	35	21.3	16	22 AAY72736	EGF-like domain #5
17	34	20.7	21	19 AAW67045	Laminin peptide an
18	34	20.7	21	19 AAW67046	Laminin peptide an
19	34	20.7	22	15 AAR53473	CD44 peptide CD44-
20	33	20.1	12	19 AAW67060	Laminin peptide an
21	33	20.1	16	19 AAW67059	Laminin peptide an
22	33	20.1	19	21 AAY98294	Alpha D peptide de
23	33	20.1	20	16 AAR72579	ENAPII active pep
24	33	20.1	21	14 AAR44441	Laminin A chain de
25	33	20.1	21	16 AAR83642	Laminin globular d
26	33	20.1	21	19 AAW67042	Laminin peptide R3
27	33	20.1	21	19 AAW67047	Laminin peptide an
28	33	20.1	21	21 AAY78577	Laminin alpha chai
29	33	20.1	21	22 AAB89238	HIV gp120 protein-
30	33	20.1	24	22 AAB88826	Human interleukin-
31	33	20.1	27	22 AAM31626	Peptide #5663 enco
32	32.5	19.8	25	20 AAY11616	Human 5' EST sece
33	32	19.5	8	22 AAB20363	Mouse prostate can
34	32	19.5	14	19 AAW67061	Laminin peptide an
35	32	19.5	16	20 AAY07887	Human secreted pro
36	32	19.5	19	21 AAY98296	Alpha D peptide de
37	32	19.5	21	19 AAW67050	Laminin peptide an
38	32	19.5	29	22 AAB48730	Mouse liver growth
39	32	19.5	30	16 AAR77215	Mouse anti-human I
40	31.5	19.2	26	13 AAR25295	Sleep inducing pep
41	31	18.9	14	19 AAW72424	Peptide #3 for red
42	31	18.9	14	21 AAY97840	Staphylococcal/str
43	31	18.9	14	22 AAB86339	G. suboxydans DSM
44	31	18.9	14	22 AAY72178	Peptide #6344, der
45	31	18.9	17	9 AAP81297	Atrial natriuretic

#### ALIGNMENTS

#### RESULT 1

AAR53482  
ID AAR53482 standard; peptide: 31 AA.  
AC AAR53482;  
XX  
DT 01-DEC-1994 (first entry)  
XX  
DE CD44 peptide CD44-9.  
XX  
KW Cell adhesion molecule; CD44; antinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autoimmune disease;  
KW immunosuppressive.  
XX  
OS Homo sapiens.  
XX  
PN WO9409811-A.  
XX  
PD 11-MAY-1994.  
XX  
PF 29-OCT-1993; 93WO-US10412.  
XX  
PR 30-OCT-1992; 92US-0973339.  
XX  
PA (UYDU-) UNIV DUKE.  
XX  
PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MU;  
XX  
DR WPI; 1994-167121/20.  
XX  
PT Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis  
XX





DR WPI: 1998-130685/12.  
DR N-PSDB; AAV15714.

XX Nucleic acid encoding secreted protein from human testis or brain  
PT cells - and related vectors, transformed cells and proteins,  
PT potentially useful e.g. as immuno-modulators, antitumour agents,  
PT promoters of tissue growth, haemostatic and thrombolytic agents etc.

XX Claim 21; Page 45; 55pp; English.

XX Clone CC397-19 has been isolated from an adult human brain cDNA library,  
CC along with clones BV2325-17, CC25-17 and BV239-3. These clones and clone  
CC BU89-13 isolated from an adult human testis library, can be expressed in  
CC eukaryotic or prokaryotic host cells. The genes and protein sequences  
CC may be used for identification purposes for example for chromosome  
CC identification or to detect genetic disorders. They can also be used to  
CC screen compounds for biological activity; to raise antibodies; as tissue  
CC markers; for isolation of related receptors and ligands and as  
CC nutritional sources. They may also have many biological activities such  
CC as cytokine and cell proliferation/differentiation activities  
CC immunosuppressant/immunostimulant activity and regulation of  
CC hematopoiesis. Neutralising antibodies against the protein can be used  
CC therapeutically.

XX Sequence 27 AA;

Query Match 23.2%; Score 38; DB 19; Length 27;  
Best Local Similarity 41.2%; Pred. No. 68;  
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LCLVPLSLAQIDLNTC 17  
||| | : | : |  
Db 7 lcipflemltislhnpvc 23

RESULT 7

AAW42020  
ID AAW42020 standard; Protein; 27 AA.

XX AC AAW42020;

DT 28-MAY-1998 (first entry)

XX Clone CC397-11 protein.

XX CC397-11 protein; human adult brain; cDNA library; cell proliferation;  
KW cell differentiation.

XX Homo sapiens.

XX US5708157-A.

XX 13-JAN-1998.

XX PF 26-JUL-1996; 96US-0686878.

XX PR 26-JUL-1996; 96US-0686878.

XX (GEMY ) GENETICS INST INC.

PA (JACO/) JACOBS K.

XX Evans C, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;

XX WPI: 1998-109352/10.

DR N-PSDB; AAV09102.

XX Human testes derived cDNA and protein(s) - may be useful for  
PT laboratory use, or for stimulating cell proliferation and  
PT differentiation

XX Disclosure; Column 57-58; 48pp; English.

XX  
CC  
CC  
CC  
CC  
XX  
SQ

This sequence is the CC397\_11 protein, which is 27 amino acids in length,  
and is isolated from adult human brain cDNA library. When the CC397\_11  
sequence was searched in the Genbank database, the clone demonstrated no  
homology with any of the known sequences.

Sequence 27 AA;

Query Match 23.2%; Score 38; DB 19; Length 27;  
Best Local Similarity 41.2%; Pred. No. 68;  
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LCLVPLSLAQIDLNTC 17  
||| | : | : |  
Db 7 lcipflemltislhnpvc 23

RESULT 8

AAV08628

ID AAV08628 standard; Protein; 30 AA.

XX AC AAV08628;

XX DT 10-AUG-1999 (first entry)

XX Human secreted protein CC397\_19.

XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;  
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;  
KW cell proliferation; cell differentiation; suppressor; tumour inhibition;  
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;  
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;  
KW cadherin; tumour invasion suppressor; gene therapy; tissue growth.

XX Homo sapiens.

XX WO9926972-A1.

XX 03-JUN-1999.

XX PF 17-NOV-1998; 98WO-US24614.

XX PR 20-OCT-1998; 98US-0175928.

PR 21-NOV-1997; 97US-0976110.

PR 18-MAY-1998; 98US-0080478.

XX (GEMY ) GENETICS INST INC.

XX Collins-Racie LA, Evans C, Jacobs K, Lavallie ER;

PI McCoy JM, Merberg D, Treacy M;

XX WPI: 1999-357813/30.

DR N-PSDB; AAX77534.

XX New polynucleotides encoding secreted proteins

PS Claim 32a; Page 113; 142pp; English.

XX This invention describes novel human secreted proteins encoded by  
CC polynucleotides isolated from human adult testes, adult brain, adult  
CC blood or adult placenta, or murine adult bone marrow or thymus cDNA  
CC libraries. The products of the invention are predicted to have biological  
CC activities which would make them suitable for treating, preventing or  
CC ameliorating medical conditions in humans and animals, although no  
CC supporting data is given. Suggested activities include nutritional  
CC activity, cytokine and cell proliferation/differentiation activity,  
CC immune stimulating (e.g. as vaccines) or suppressing activity,  
CC haematopoiesis regulating activity, tissue growth activity,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, cadherin/tumour invasion suppressor activity, and tumour  
CC inhibition activity. The polynucleotides are also stated to be useful

CC for gene therapy.

SQ Sequence 30 AA;

Query Match 23.2%; Score 38; DB 20; Length 30;

Best Local Similarity 41.2%; Pred. No. 77;

Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 LCLVPLSLAQIDLNITC 17

||| | : : ||: |

Db 7 lclpflmitlsinvpc 23

RESULT 9

AAB34787

ID AAB34787 standard; Protein; 30 AA.

XX AC AAB34787;

XX DT 26-JAN-2001 (first entry)

XX DE Human secreted protein sequence encoded by gene 15 SEQ ID NO:75.

XX DE Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;  
KW cancer; immune disorder; cardiovascular disorder; wound healing;  
KW neurological disease; infectious disease; chromosome identification.

OS Homo sapiens.

XX WO200058356-A1.

XX PD 05-OCT-2000.

XX XX 22-MAR-2000; 2000WO-US07535.

XX PF 26-MAR-1999; 99US-0126511.

XX PR 17-DEC-1999; 99US-0172413.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX DR WPI; 2000-594639/56.

XX DR N-PSDB; AAC59980.

XX PT Fifty nucleic acid molecules encoding human secreted proteins, useful  
in the prevention, treatment and diagnosis of cancer, immune disorders,  
PT cardiovascular disorders and neurological diseases -

XX PS Claim 1; Page 381; 425pp; English.

XX CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the  
human secreted proteins given in AAB34773 to AAB34822. AAB34823 to  
AAB34852 represent human secreted polypeptide sequences and proteins  
homologous to them, which are given in the exemplification of the present  
invention. Human secreted proteins have activities based on the tissues  
and cells the genes are expressed in. Examples of activities include:  
CC cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;  
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and  
CC cardiant. The polynucleotides and polypeptides are useful for  
CC preventing, treating or ameliorating a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. The polynucleotides are  
CC useful for chromosome identification. They are also useful as probes for  
CC diagnosing a disorder related to the female reproductive system,  
CC particularly breast and/or ovary cancer. They are also useful in the gene  
therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,

CC agonists and antagonists from the present invention are useful in the  
CC diagnosis, treatment and prevention of cancer, immune disorders,  
CC cardiovascular disorders, wound healing, neurological diseases and  
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence  
CC used in the exemplification of the present invention.

XX SQ Sequence 30 AA;

Query Match 23.2%; Score 38; DB 21; Length 30;

Best Local Similarity 63.6%; Pred. No. 77;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LCLVPLSLAQI 11

:||||||| |:

Db 16 mclvplspahv 26

RESULT 10

AAY67319

ID AAY67319 standard; Protein; 30 AA.

XX AC AAY67319;

XX DT 11-APR-2000 (first entry)

XX DE Human secreted protein CC397\_19 amino acid sequence.

XX DE Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;  
KW placental pathology; metastasis inhibition; nutritional activity;  
KW immune stimulator; haematopoiesis regulator; tissue growth;  
KW tumour inhibitor; anti-inflammatory; clone CC397\_19; ATCC-98153.

OS Homo sapiens.

XX WO9960020-A1.

XX PD 25-NOV-1999.

XX PF 17-MAY-1999; 99WO-US10915.

XX PR 18-MAY-1998; 98US-0080478.

XX PR 20-OCT-1998; 98US-0175928.

XX PA (GEMY) GENETICS INST INC.

XX PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX PI Merberg D, Mi S, Treacy M;

XX DR WPI; 2000-116311/10.

XX DR N-PSDB; AAZ59476.

XX PT New polynucleotides encoding secreted cDNA libraries, used to develop  
products for the diagnosis and treatment of neoplastic disease .

XX PS Claim 28; Page 120; 149pp; English.

XX CC This is the human secreted protein CC397\_19 amino acid sequence,  
the polynucleotide sequence encoding CC397\_19 was obtained from a human  
adult brain cDNA library. The invention relates to secreted human and  
murine proteins. The polynucleotides and proteins are predicted to have  
CC biological activities which would make them suitable for treating,  
preventing or ameliorating medical conditions in humans and animals.  
CC Detection of the levels of the proteins can be used for the diagnosis of  
e.g. pre-eclampsia, placental pathology or cancer. Agents which modulate  
the expression or function of the proteins may be used for treating a  
CC neoplastic disease and inhibiting metastasis. Other suggested activities  
include nutritional activity (e.g. in feeds), cytokine and cell  
proliferation/differentiation activity, immune stimulating (e.g. as  
vaccines) or suppressing activity, haematopoiesis regulating activity,  
tissue growth activity, activin/inhibin activity,  
chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
receptor/ligand activity, anti-inflammatory activity, cadherin/tumour

CC invasion suppressor activity, and tumour inhibition activity. The  
 CC polynucleotide sequences are also stated to be useful for gene therapy.  
 XX  
 SQ Sequence 30 AA;

Query Match 23.2%; Score 38; DB 21; Length 30;  
 Best Local Similarity 41.2%; Pred. No. 77;  
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 1 LCLVPLSLAQIDLNITC 17  
 ||| | : : ||: |  
 Db 7 lclpflemltislrvpc 23

RESULT 11  
 AAEO2383  
 ID AAEO2383 standard; peptide; 23 AA.

XX AC AAEO2383;  
 DT 10-AUG-2001 (first entry)  
 XX Human alpha1-antitrypsin signal peptide.

XX Human; Factor VIII:C; FVIII:C; coagulation; von-Wille-brand factor;  
 KW type A haemophilia; blood clotting disorder; haemostatic; gene therapy;  
 KW recombinant protein; immunogen; alpha1-antitrypsin.

XX OS Homo sapiens.

XX PN US6228620-B1.

XX PD 08-MAY-2001.

XX PF 16-MAY-1995; 95US-0441943.

XX PR 03-DEC-1993; 93US-0161770.

XX PR 07-FEB-1991; 91US-0652099.

XX PR 27-JAN-1986; 86US-0822989.

XX PR 19-MAY-1987; 87US-0051916.

XX PA (CHIR ) CHIRON CORP.

XX PA (NOVO ) NOVO NORDISK AS.

XX PI Chapman B, Burke RL, Rasmussen ME, Mikkelsen JM;

XX DR WPI; 2001-335014/35.

XX PT New nucleic acid complexes having factor VIII activity and coagulation  
 PT activity, for treating hemophilia and other blood clotting disorders,  
 PT or as immunogens for the production of antibodies

XX PS Example 6; Column 51-52; 33pp; English.

XX CC The patent relates to a nucleic acid composition for introducing  
 CC nucleic acid into a eukaryotic host cell to obtain expression of a  
 CC recombinant protein which lacks all or a portion of the B domain of  
 CC human Factor VIII:C (hFVIII:C) and is capable of coagulation activity.  
 CC The recombinant protein consists of a polypeptide substantially  
 CC homologous to mature A domain of hFVIII:C as encoded by the  
 CC polynucleotide present in plasmid pSVF8-200 (ATCC number 40190); and  
 CC a second polypeptide homologous to mature C domain of hFVIII:C as  
 CC encoded by the polynucleotide present in plasmid pSVF8-200.

XX CC The composition is useful as an immunogen for the production of  
 CC antibodies, for isolating von-Wille-brand factor by affinity  
 CC chromatography, in diagnostic assays for Factor VIII:C and for treating  
 CC type A haemophilia and other blood clotting disorders.  
 CC The present sequence is human alpha1-antitrypsin signal peptide  
 CC used for secretion of recombinant protein of the invention.

XX SQ Sequence 23 AA;

Query Match 22.6%; Score 37; DB 22; Length 23;  
 Best Local Similarity 87.5%; Pred. No. 80;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CLVPLSLA 9  
 ||||:||||  
 Db 16 clvpvsia 23

RESULT 12

AAEP82918  
 ID AAP82918 standard; protein; 29 AA.

XX AC AAP82918;

XX DT 23-NOV-1990 (first entry)

XX DE Product from plasmid pTG2902 for extracellular hirudin expression,  
 DE signal sequence cleavage site.

XX KW Hirudin expression vectors; signal sequence cleavage site; pTG2902;  
 KW thrombin inhibitor; anticoagulant; alpha-antitrypsin.

XX FH Key Location/Qualifiers

XX FT Peptide 1..24

XX FT /label=signal\_sequence

XX FT /note="alpha-antitrypsin"

XX FT 25..29

XX FT /label=hirudine

XX FT /note="native"

XX FT Cleavage-site 24..25

XX FT /label=StuI

XX PN FR2611723-A.

XX PD 09-SEP-1988.

XX PF 27-FEB-1987; 87FR-0002696.

XX PR 27-FEB-1987; 87FR-0002696.

XX PA (TRAN-) TRANSGENE SA.

XX PI Skern T, Courtney M;

XX DR WPI; 1988-301637/43.

XX DR P-PSDB; P829187.

XX PT Hirudin expression vectors - contg. signal sequence allowing  
 PT extracellular hirudin prodn. by eukaryotic cells.

XX PS Disclosure; 10pp; French.

XX CC The DNA sequence encoding native hirudine is located downstream from the  
 CC alpha-antitrypsin sequence ensuring the expression and release of  
 CC hirudin in active form by cells into the culture medium.

XX CC Between the two sequences is a StuI site. Hirudin is a specific  
 CC thrombin inhibitor useful as an anticoagulant.

XX CC See also AAN82293-95.

XX SQ Sequence 29 AA;

Query Match 22.6%; Score 37; DB 9; Length 29;

Best Local Similarity 87.5%; Pred. No. 1e+02;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CLVPLSLA 9

||||:||||

Db 17 clvpvsia 24



## RESULT 13

AAW67044  
ID AAW67044 standard; peptide; 21 AA.

XX AC AAW67044;

XX DT 16-DEC-1998 (first entry)

XX DE Laminin peptide analogue 5104.

XX laminin; systemic lupus erythematosus; SLE; R38; 5104; autoantibody;  
KW analogue.

XX OS Synthetic.

XX OS Musc sp.

XX PN WO9842737-A2.

XX PD 01-OCT-1998.

XX PF 20-MAR-1998; 98WO-IB00415.

XX PR 20-MAR-1997; 97IL-0120503.

XX PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

XX PI Naparstek Y;

XX DR WPI; 1998-583143/49.

XX Treatment of systemic lupus erythematosus - comprise administration  
PT of laminin peptides, particularly R38 and R38 derivatives and  
PT analogues to mammal

XX PS Claim 11; Page 19; 28pp; English.

XX The invention relates to a method for treating systemic lupus  
CC erythematosus (SLE), which comprises administering an effective dose of  
CC at least one laminin peptide, or a laminin peptide analogue or  
CC derivative. In one exemplary embodiment, the laminin peptide is selected  
CC from the group consisting of R38 and claimed R38 analogues and  
CC derivatives including 5200, 5104, 5105, 5106, 5107, 5108, 5109 and 5110.  
CC SLE is serologically characterised by the presence of autoantibodies in  
CC the serum, particularly anti-DNA autoantibodies, and the methods and  
CC products may be used to treat SLE using laminin derived peptides which  
CC cross-react with those antibodies. At least one laminin peptide may be  
CC used to prepare a therapeutic for administration to a mammal suffering  
CC from SLE. The treatment is not associated with the side effects of prior  
CC art immunosuppressive treatment, and the monitoring assay is more  
CC specific and less invasive than prior art assays. The present sequence  
CC represents the laminin peptide analogue 5104.

XX SQ Sequence 21 AA;

Query Match 22.0%; Score 36; DB 19; Length 21;

Best Local Similarity 70.0%; Pred. No. 1e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 IDLNTCRFA 20

Db 8 ldlntlefa 17

## RESULT 14

AAAB84641

ID AAB84641 standard; peptide; 26 AA.

XX AC AAB84641;

XX DT 05-SEP-2001 (first entry)

XX DE Antigenic peptide derived from the MNKD protein.

XX

KW Dishevelled-associated protein 1A; DAP1A; mnkd; dishevelled protein;  
KW Wnt signalling; colon cancer; cancer.

XX OS Mus sp.

XX PN WO200144279-A2.

XX PD 21-JUN-2001.

XX PF 05-DEC-2000; 2000WO-US32986.

XX PR 17-DEC-1999; 99US-0172434.

XX PA (CHIR ) CHIRON CORP.

XX PI Yan D, Williams LT;

XX DR WPI; 2001-425440/45.

XX Novel mammalian dishevelled-associated proteins, mnkd and DAP1A, useful  
PT for inhibiting Wnt signalling in mammalian cells, and thus for treating  
PT colon cancer

XX PS Disclosure; Page 34; 67pp; English.

XX The present sequence represents an antigenic peptide of mnkd. The  
CC peptide was used to raise antibodies. The specification describes  
CC mnkd and dishevelled-associated protein 1A (DAP1A) proteins, that  
CC interact with mammalian dishevelled protein. mnkd is useful for  
CC inhibiting Wnt signalling in mammalian cells, and thus for treating  
CC colon cancer. mnkd is also useful for activating the JNK pathway.  
CC mnkd and DAP1A are also useful for screening drugs that are useful  
CC for treating cancer.

XX SQ Sequence 26 AA;

Query Match 22.0%; Score 36; DB 22; Length 26;

Best Local Similarity 46.7%; Pred. No. 1.3e+02;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 17 CRFAGVFHVERNGRY 31

Db 1 crfggdshleqdcy 15

## RESULT 15

AAAY19691

ID AAY19691 standard; Protein; 28 AA.

XX AC AAY19691;

XX DT 14-JUL-1999 (first entry)

XX SEQ ID NO 409 from WO9922243.

XX Human secreted protein; cancer; tumour; neurodegenerative disorder;  
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;  
KW immune system disease; autoimmune disease; hepatic disease;  
KW renal disease; lymphoma; inflammation; allergy; ischemic shock;  
KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;  
KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;  
KW lung disease; thymus disease; digestive disorder; endocrine disorder;  
KW infection; AIDS.

XX OS Homo sapiens.

XX PN WO9922243-A1.

XX PD 06-MAY-1999.

XX PF 23-OCT-1998; 98WO-US22376.

XX 24-OCT-1997; 97US-0063387.  
PR 24-OCT-1997; 97US-0082784.  
PR 24-OCT-1997; 97US-0063088.  
PR 24-OCT-1997; 97US-0063089.  
PR 24-OCT-1997; 97US-0063090.  
PR 24-OCT-1997; 97US-0063091.  
PR 24-OCT-1997; 97US-0063092.  
PR 24-OCT-1997; 97US-0063097.  
PR 24-OCT-1997; 97US-0063098.  
PR 24-OCT-1997; 97US-0063099.  
PR 24-OCT-1997; 97US-0063100.  
PR 24-OCT-1997; 97US-0063101.  
PR 24-OCT-1997; 97US-0063109.  
PR 24-OCT-1997; 97US-0063110.  
PR 24-OCT-1997; 97US-0063111.  
PR 24-OCT-1997; 97US-0063148.  
PR 24-OCT-1997; 97US-0063386.  
XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX  
PI Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;  
PI Feng P, Florence C, Florence KA, Greene JM, Janat F;  
PI Kayw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;  
PI Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;  
XX

DR WPI; 1999-303069/25.

DR  
XX New isolated human genes and the secreted polypeptides they encode

PT  
XX Disclosure; Page 499; 546pp; English.

XX  
CC The specification describes human secreted proteins. The polynucleotides  
CC and their corresponding secreted polypeptides are useful for preventing,  
CC treating or ameliorating medical conditions, e.g. by protein or gene  
CC therapy. Pathological conditions can also be diagnosed by determining  
CC the amount of the polypeptides in a sample or by determining the presence  
CC of mutations in the polynucleotides. Specific uses are described for each  
CC of the polynucleotides, based on which tissues they are most highly  
CC expressed in, and include developing products for the diagnosis or  
CC treatment of cancer, tumours, neurodegenerative disorders, developmental  
CC abnormalities and fetal deficiencies, blood disorders, leukemias,  
CC diseases of the immune system, autoimmune diseases, hepatic and renal  
CC disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's  
CC and cognitive disorders, schizophrenia, prostate diseases, obesity,  
CC disorders involving osteoclasts such as osteoporosis, arthritis or  
CC malignancies, diseases of testes, lung or thymus, digestive/endocrine  
CC disorders, infections and AIDS. The polypeptides are also useful for  
CC identifying their binding partners.  
XX

SQ Sequence 28 AA;

Query Match 22.0%; Score 36; DB 20; Length 28;  
Best Local Similarity 26.3%; pred. No. 1.4e+02;  
Matches 5; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 LCLVPLSLAQIDLNITCRF 19  
| | | : : : | : |  
Db 4 lcimkivipvmsmktncqf 22

Search completed: February 21, 2002, 17:08:22  
Job time: 188 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:06:49 ; Search time 12.47 Seconds  
(without alignments)  
55.942 Million cell updates/sec

Title: US-08-753-851-12  
Perfect score: 164  
Sequence: 1 LCLVPLSLAQIDLNITCRFAGVHVEKNGRY 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 133741

Minimum DB seq length: 0  
Maximum DB seq length: 31

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	164	100.0	31	2	US-08-143-311B-10
2	69	42.1	25	2	US-08-143-311B-27
3	58	35.4	25	2	US-08-143-311B-26
4	39	23.8	25	1	US-07-679-052A-12
5	38	23.2	27	1	US-08-686-878A-17
6	37.5	22.9	27	2	US-08-620-151-105
7	37	22.6	23	4	US-08-441-943-35
8	37	22.6	24	1	US-07-679-052A-2
9	36	22.0	21	4	US-09-399-494-15
10	35	21.3	20	4	US-08-360-821B-39
11	35	21.3	23	2	US-08-637-759B-499
12	35	21.3	23	3	US-08-871-355A-499
13	34	20.7	21	4	US-09-399-494-16
14	34	20.7	21	4	US-09-399-494-17
15	34	20.7	22	2	US-08-143-311B-1
16	34	20.7	23	2	US-08-687-702-5
17	34	20.7	26	2	US-08-467-587A-14
18	33	20.1	12	4	US-09-399-494-13
19	33	20.1	16	4	US-09-399-494-12
20	33	20.1	20	1	US-08-129-456A-2
21	33	20.1	20	4	US-08-360-821B-2
22	33	20.1	21	3	US-08-394-748A-6
23	33	20.1	21	4	US-09-399-494-1
24	33	20.1	21	4	US-09-399-494-7
25	33	20.1	21	4	US-09-399-494-9
26	33	20.1	21	4	US-09-399-494-18
27	33	20.1	21	5	PCT-US95-02478-6

28 33 20.1 25 1 US-07-961-724C-2 Sequence 2, Appli  
29 33 20.1 26 2 US-08-467-587A-21 Sequence 21, Appli  
30 33 20.1 27 2 US-08-846-526-2 Sequence 2, Appli  
31 32 19.5 14 4 US-09-399-494-14 Sequence 14, Appli  
32 32 19.5 21 4 US-09-399-494-21 Sequence 21, Appli  
33 32 19.5 30 2 US-08-553-501A-76 Sequence 76, Appli  
34 32 19.5 30 2 US-08-553-501A-80 Sequence 80, Appli  
35 32 19.5 30 3 US-09-205-231-76 Sequence 76, Appli  
36 32 19.5 30 3 US-09-205-231-80 Sequence 80, Appli  
37 31.5 19.2 15 1 US-08-462-949-17 Sequence 17, Appli  
38 31.5 19.2 15 1 US-08-023-764B-17 Sequence 17, Appli  
39 31 18.9 14 3 US-08-838-413A-5 Sequence 5, Appli  
40 31 18.9 20 5 PCT-US94-10257A-77 Sequence 77, Appli  
41 31 18.9 22 4 US-09-240-078-23 Sequence 23, Appli  
42 31 18.9 31 1 US-08-440-103-28 Sequence 28, Appli  
43 31 18.9 31 1 US-08-440-542-28 Sequence 28, Appli  
44 31 18.9 31 1 US-08-231-368-28 Sequence 28, Appli  
45 31 18.9 31 1 US-08-440-210-28 Sequence 28, Appli

## ALIGNMENTS

RESULT 1  
US-08-143-311B-10  
; Sequence 10, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-10



; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/679,052A  
; FILING DATE: 19910506  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16786/147 CHAC  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-679-052A-12

Query Match 23.8%; Score 39; DB 1; Length 25;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CLVPLSLAQ 10  
Db 17 CLVPVSLAE 25

RESULT 5  
US-08-686-878A-17  
; Sequence 17, Application US/08686878A  
; Patent No. 5708157  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John  
; APPLICANT: Lavallie, Edward  
; APPLICANT: Racie, Lisa  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Spaulding, Vikki  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 Bridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/686,878A  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-686-878A-17

Query Match 23.2%; Score 38; DB 1; Length 27;  
Best Local Similarity 41.2%; Pred. No. 17;  
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LCLVPLSLAQIDLNITC 17  
Db 7 LCLPFLEMLTSLNVP 23

RESULT 6  
US-08-620-151-105  
; Sequence 105, Application US/08620151  
; Patent No. 5928955  
; GENERAL INFORMATION:  
; APPLICANT: Imperiali, Barbara  
; APPLICANT: Walkup, Grant K.  
; TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR  
; TITLE OF INVENTION: DIVALENT ZINC  
; NUMBER OF SEQUENCES: 136  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
; STREET: Plaza Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60611-5599  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620,151  
; FILING DATE: 22-MAR-1996  
; CLASSIFICATION: 422  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shannon, Karen L.  
; REGISTRATION NUMBER: 36,675  
; REFERENCE/DOCKET NUMBER: 8597/6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-321-4200  
; TELEFAX: 312-321-4299  
; INFORMATION FOR SEQ ID NO: 105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-620-151-105

Query Match 22.9%; Score 37.5; DB 2; Length 27;  
Best Local Similarity 34.6%; Pred. No. 21;  
Matches 9; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

QY 5 PLSLAQIDLNITCREFAGVFH-VEKNG 29  
Db 1 PTKCTCDATYQCRSSAVTHMYNKHG 26

RESULT 7  
 US-08-441-943-35  
 ; Sequence 35, Application US/08441943  
 ; Patent No. 6228620  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHAPMAN, BARBARA  
 ; APPLICANT: BURKE, RAE LYNN  
 ; APPLICANT: RASMUSSEN, MIRELLA EZBAN  
 ; APPLICANT: MIKKELSON, JAN MOLLER  
 ; TITLE OF INVENTION: PROTEIN COMPLEXES HAVING FACTOR VIII:C  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: REED & ROBINS  
 ; STREET: 285 HAMILTON AVENUE, SUITE 200  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: UNITED STATES OF AMERICA  
 ; ZIP: 94301  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/441,943  
 ; FILING DATE: 16-MAY-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 051,916  
 ; FILING DATE: 19-MAY-1987  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 822,989  
 ; FILING DATE: 27-JAN-1986  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BAROVSKI, KENNETH  
 ; REGISTRATION NUMBER: 36,442  
 ; REFERENCE/DOCKET NUMBER: 2300-0048.10  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 327-3400  
 ; TELEFAX: (415) 327-3231  
 ; INFORMATION FOR SEQ ID NO: 35:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-441-943-35

Query Match 22.6%; Score 37; DB 4; Length 23;  
 Best Local Similarity 87.5%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CLVPLSLA 9  
 |||||  
 Db 16 CLVPSVSLA 23

RESULT 8  
 US-07-679-052A-2  
 ; Sequence 2, Application US/07679052A  
 ; Patent No. 5258400  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WHITEFELD, Peter L.  
 ; APPLICANT: RICHARDSON, Michael A.  
 ; APPLICANT: BUNN, Clive L.  
 ; TITLE OF INVENTION: RECOMBINANT PRODUCT  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/679,052A  
 ; FILING DATE: 19910506  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 16786/147 CHAC  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703)836-9300  
 ; TELEFAX: (703)683-4109  
 ; TELEX: 899149  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 24 amino acids  
 ; TYPE: AMINO ACID  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 ; FEATURE:  
 ; NAME/KEY: Peptide  
 ; LOCATION: 1..24  
 ; OTHER INFORMATION: /note= "Signal peptide from human  
 ; OTHER INFORMATION: a-1-antitrypsin"  
 ; US-07-679-052A-2

Query Match 22.6%; Score 37; DB 1; Length 24;  
 Best Local Similarity 87.5%; Pred. No. 21;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CLVPLSLA 9  
 |||||  
 Db 17 CLVPSVSLA 24

RESULT 9  
 US-09-399-494-15  
 ; Sequence 15, Application US/09399494  
 ; Patent No. 6228363  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Naparstek, Yaakov  
 ; TITLE OF INVENTION: PEPTIDES FOR THE TREATMENT OF SYSTEMIC LUPUS  
 ; FILE REFERENCE: 56040-A-PCT-US  
 ; CURRENT APPLICATION NUMBER: US/09/399,494  
 ; CURRENT FILING DATE: 1999-09-20  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 15  
 ; LENGTH: 21  
 ; TYPE: PPT  
 ; ORGANISM: mouse  
 ; US-09-399-494-15

Query Match 22.0%; Score 36; DB 4; Length 21;  
 Best Local Similarity 70.0%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 11 IDLNITCRFA 20  
 :|||||  
 Db 8 LDLNITLFA 17

```
RESULT 10
US-08-360-821B-39
; Sequence 39, Application US/08360821B
; Patent No. 6228837
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Clauss, Matthias
; APPLICANT: Kao, Janet
; APPLICANT: Kayton, Mark
; APPLICANT: Libutti, Steven K
; TITLE OF INVENTION: Endothelial Monocyte Activating
; TITLE OF INVENTION: Polypeptide II: A Mediator Which Activates Host Response
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.30, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICANT: PatentIn Release #1.30, Version #1.30
; FILING DATE: 08-OCT-96
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-360-821B-39

Query Match 21.3%; Score 35; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 35;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 PLSLAQIDLNITC 17
Db 2 PIDVSRDLRIGC 14

RESULT 11
US-08-637-759B-499
; Sequence 499, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 499:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-637-759B-499

Query Match 21.3%; Score 35; DB 2; Length 23;
Best Local Similarity 45.5%; Pred. No. 41;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 21 GVFHVKNRGY. 31
Db 8 GLYHLRKDHY 18

RESULT 12
US-08-871-355A-499
; Sequence 499, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
```

; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 499:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-871-355A-499

Query Match 21.3%; Score 35; DB 3; Length 23;  
Best Local Similarity 45.5%; Pred. No. 41;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 21 GVHVXKNGRY 31  
I::I::I::I  
Db 8 GLYHLRKDHRY 18

RESULT 13  
US-09-399-494-16  
; Sequence 16, Application US/09399494  
; Patent No. 6228363  
; GENERAL INFORMATION:  
; APPLICANT: Naparstek, Yaakov  
; TITLE OF INVENTION: PEPTIDES FOR THE TREATMENT OF SYSTEMIC LUPUS  
; FILE REFERENCE: 56040-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/399,494  
; CURRENT FILING DATE: 1999-09-20  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: mouse  
US-09-399-494-16

Query Match 20.7%; Score 34; DB 4; Length 21;  
Best Local Similarity 60.0%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIDLNITCRF 19  
:::|||||I  
Db 7 ELDLNTLEF 16

RESULT 14  
US-09-399-494-17  
; Sequence 17, Application US/09399494  
; Patent No. 6228363  
; GENERAL INFORMATION:  
; APPLICANT: Naparstek, Yaakov  
; TITLE OF INVENTION: PEPTIDES FOR THE TREATMENT OF SYSTEMIC LUPUS  
; FILE REFERENCE: 56040-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/399,494  
; CURRENT FILING DATE: 1999-09-20  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: mouse  
US-09-399-494-17

Query Match 20.7%; Score 34; DB 4; Length 21;  
Best Local Similarity 60.0%; Pred. No. 53;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 10 QIDLNITCRF 19  
:::|||||I  
Db 7 ELDLNTLEF 16

RESULT 15  
US-08-143-311B-1  
; Sequence 1, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARIJLYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-1

Query Match 20.7%; Score 34; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 EKNGRY 31  
|||||I  
Db 2 EKNGRY 7

Search completed: February 21, 2002, 17:08:41  
Job time: 112 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:11:04 ; Search time 12.69 seconds  
(without alignments)  
168.076 Million cell updates/sec

Title: US-08-753-851-13  
Perfect score: 147  
Sequence: 1 LKAFNSTLPTMAQMEKALSIGFTCRY 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 5648

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33.5	22.8	16	2 S33590	beta-crystallin A3
2	30	20.4	19	2 S67517	Na+/K+-exchanging
3	30	20.4	22	2 B60771	CDw40 antigen Hu54
4	29	19.7	24	1 B32252	pyrroloquinoline q
5	29	19.7	24	1 S82422	pyrroloquinoline q
6	27.5	18.7	19	2 S68394	H+-transporting AT
7	27	18.4	18	2 A54651	insulin-like growt
8	27	18.4	20	2 A60802	25K acrosomal auto
9	27	18.4	28	2 A32691	apolipoprotein C-I
10	26	17.7	16	2 D32395	ribulose-bisphosph
11	26	17.7	21	2 A85712	unknown protein en
12	26	17.7	23	2 S60567	homeodomain protei
13	26	17.7	23	2 A83397	pyrroloquinoline q
14	26	17.7	24	2 A53357	neurotoxin Bt-II-
15	26	17.7	25	2 A32203	dihydrofolate redu
16	26	17.7	27	2 I44629	homeotic protein H
17	25.5	17.3	15	2 PA0020	protein QA100028 -
18	25.5	17.3	26	4 A01640	hypothetical prote
19	25.5	17.3	27	2 T17064	NADH dehydrogenase
20	25	17.0	17	2 D47274	calcium-binding pr
21	25	17.0	27	2 I57644	transmembrane gly
22	25	17.0	24	2 F38746	pro-B cell differe
23	24.5	16.7	20	2 A42865	Ca2+/calmodulin-de
24	24	16.3	13	2 S36668	hypothetical prote
25	24	16.3	13	2 A61288	spore proteinase g
26	24	16.3	13	2 B58533	CD61 homolog - cha
27	24	16.3	13	2 B19434	probable sex-speci
28	24	16.3	22	2 S07817	spike glycoprotein
29	24	16.3	22	2 T01859	CYI 1a protein - c

30 24 16.3 25 2 B69274  
31 24 16.3 27 2 F44629  
32 24 16.3 27 2 B44636  
33 23.5 16.0 19 2 B39845  
34 23.5 16.0 22 2 PT0316  
35 23 15.6 17 2 A27636  
36 23 15.6 18 2 A41877  
37 23 15.6 18 2 S45373  
38 23 15.6 19 2 PC2326  
39 23 15.6 19 2 B60822  
40 23 15.6 21 2 A20359  
41 23 15.6 23 2 I45916  
42 23 15.6 23 2 H45087  
43 23 15.6 26 2 S07303  
44 23 15.6 27 2 B61318  
45 23 15.6 27 4 S20754

hypothetical prote  
homeotic protein H  
homeotic protein H  
pyrB leader peptid  
Ig heavy chain CDR  
cytotoxin B - C1os  
LcrKc - Versinia p  
translation elonga  
multicatalytic end  
cytochrome P450 U  
translation elonga  
epidermal keratin  
cysteine proteinase  
ribulose-bisphosph  
momordin - balsam  
probable pre-core

## ALIGNMENTS

RESULT 1

S33590  
beta-crystallin A3 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999  
C:Accession: S33590  
R:David, L.L.; Shearer, T.R.  
FEBS Lett. 324, 265-270, 1993  
A:Title: beta-crystallins insolubilized by calpain II in vitro contain cleavage sites  
A:Reference number: S33586; MUID:94009594  
A:Accession: S33590  
A:Molecule type: protein  
A:Residues: 1-16 <DAV>

Query Match 22.8%; Score 33.5; DB 2; Length 16;  
Best Local Similarity 47.1%; Pred. No. 83;  
Matches 8; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 8 TLPTMAQMEKALSIGFE 24  
|||||::|  
DB 1 TLPT-TKVEGXGAXIGYE 16

RESULT 2

S67517  
Na+/K+-exchanging ATPase (EC 3.6.1.37) alpha chain - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 17-Mar-1999  
C:Accession: S67517  
R:Antolovic, R.; Linder, D.; Hahnen, J.; Schoner, W.  
Eur. J. Biochem. 227, 61-67, 1995  
A:Title: Affinity labeling of a sulphydryl group in the cardiacylgcoside receptor sit  
A:Reference number: S67517; MUID:95154347  
A:Accession: S67517  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-19 <ANT>  
C:Keywords: hydrolase

Query Match 20.4%; Score 30; DB 2; Length 19;  
Best Local Similarity 31.2%; Pred. No. 3.5e+02;  
Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSTLPTMAQMEKALSI 21  
|::|::|:  
DB 4 NALTPDMDLAKKXVSM 19

RESULT 3

B60771

CDw40 antigen Hu549 - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 26-May-2000  
C:Accession: B60771  
R:Brasch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.  
J. Immunol. 142, 562-567, 1989  
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-like  
A:Reference number: A60771; MUID:89093941

A:Accession: B60771  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <BRA>  
C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 20.4%; Score 30; DB 1; Length 22;  
Best Local Similarity 41.2%; Pred. No. 4e+02;  
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 10 PTMAQMEKALSIGFETC 26  
I I I I I I I I  
DB 2 PPTAQREKQYLINGQCC 18

RESULT 4

B32252  
pyrroloquinoline quinone precursor - Acinetobacter calcoaceticus  
N:Alternate names: pqq gene IV protein  
C:Species: Acinetobacter calcoaceticus  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C:Accession: B32252  
R:Goosen, N.; Horsman, H.P.A.; Huinen, R.G.M.; van de Putte, P.  
J. Bacteriol. 171, 447-455, 1989  
A:Title: Acinetobacter calcoaceticus genes involved in biosynthesis of the coenzyme pyrroloquinoline quinone  
A:Reference number: A32252; MUID:89123056

A:Accession: B32252  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-24 <GOO>  
A:Cross-references: GB:X06452; NID:g38740; PID:g130796  
A:Note: this ORF is not annotated in Genbank entry ACPQSYN, release 116.0  
R:Goosen, N.; Huinen, R.G.; van de Putte, P.  
J. Bacteriol. 174, 1426-1427, 1992  
A:Title: A 24-amino-acid polypeptide is essential for the biosynthesis of the coenzyme pyrroloquinoline quinone  
A:Reference number: A59183; MUID:92138642  
A:Contents: annotation  
C:Genetics:

A:Gene: pqqA  
C:Superfamily: pyrroloquinoline quinone precursor pqqA  
C:Keywords: quinoprotein

F:16-20/Product: pyrroloquinoline quinone #status predicted <MAT>  
F:16-20/Cross-link: Pyrroloquinoline quinone (Glu, Tyr) #status predicted

Query Match 19.7%; Score 29; DB 1; Length 24;  
Best Local Similarity 60.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 19 LSGIFETCRY 28  
I I I I I I  
DB 11 LRIGFEVTMY 20

RESULT 5

S58242  
pyrroloquinoline quinone precursor pqqA - Pseudomonas fluorescens  
N:Alternate names: pyrroloquinoline quinone biosynthesis A  
C:Species: Pseudomonas fluorescens  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
C:Accession: S58242  
R:Schlader, U.; Keel, C.; Defago, G.; Haas, D.  
submitted to the EMBL Data Library, May 1995  
A:Description: Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHA0: the

A:Reference number: S58239

A:Accession: S58242

A:Molecule type: DNA

A:Residues: 'MYRQHPSHPORSNF', 1-24 <SCH>

A:Cross-references: EMBL:X87299; NID:g929799; PIDN:CAA60731.1; PID:g929803; PIDN:CAA6

A:Note: in Genbank entry PFQOABCF, release 116.0, the indicated alternative sequence

C:Genetics:

A:Gene: pqqA

C:Superfamily: pyrroloquinoline quinone precursor pqqA

C:Keywords: quinoprotein

F:16-20/Product: pyrroloquinoline quinone #status predicted <MAT>

F:16-20/Cross-link: pyrroloquinoline quinone (Glu, Tyr) #status predicted

Query Match 19.7%; Score 29; DB 1; Length 24;  
Best Local Similarity 60.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 19 LSGIFETCRY 28  
I I I I I I I I  
DB 11 LRIGFEVTMY 20

RESULT 6

S68394  
H+-transporting ATP synthase (EC 3.6.1.34) chain G - Chlamydomonas reinhardtii (fragm  
N:Alternate names: ATP synthase chain G  
C:Species: Chlamydomonas reinhardtii  
C:Date: 04-Dec-1997 #sequence\_revision 12-Dec-1997 #text\_change 17-Mar-1999  
C:Accession: S68394  
R:Fiedler, H.R.; Schmid, R.; Leu, S.; Shavit, N.; Strotmann, H.  
FEBS Lett. 377, 163-166, 1995  
A:Title: Isolation of CF(0)CF(1) from Chlamydomonas reinhardtii cw15 and the N-termin  
A:Reference number: S68388; MUID:96128220

A:Accession: S68394  
A:Molecule type: protein

A:Residues: 1-19 <FIE>

A:Experimental source: strain CW15

C:Genetics:

A:Genome: nuclear

C:Keywords: chloroplast; hydrolase

Query Match 18.7%; Score 27.5; DB 2; Length 19;  
Best Local Similarity 53.3%; Pred. No. 8.4e+02;  
Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 3 KAFN-STLPTMAQME 16  
I I I I I I I I  
DB 4 KIFDNTLPVMAGEE 18

RESULT 7

A54651  
insulin-like growth factor-binding protein 22K form - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 13-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 11-Apr-1995  
C:Accession: A54651  
R:Roghani, M.; Segovia, B.; Whitechurch, O.; Binoux, M.

Growth Regul. 1, 125-130, 1991

A:Title: Purification from human cerebrospinal fluid of insulin-like growth factor bi

A:Reference number: A54651; MUID:93091816

A:Accession: A54651

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-18 <ROG>

A:Experimental source: cerebrospinal fluid

A:Note: sequence extracted from NCBI backbone (NCBIP:121074)

Query Match 18.4%; Score 27; DB 2; Length 18;

Best Local Similarity 46.7%; Pred. No. 9.4e+02;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

[illegible]

Db 1 FNKYLTPARRLEIAAALG 18

## RESULT 13

A83397  
pyroloquinoline quinone biosynthesis protein A PAL985 [imported] - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: A83397  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: A82950; MUID:20437337  
A:Accession: A83397  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-23 <STO>  
A:Cross-references: GB:AE004625; GB:AE004091; MID:g9947983; PIDN:AAG05373.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: pqqA; PAL985

Query Match 17.7%; Score 26; DB 2; Length 23;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 19 LSIGFETCRY 28  
| :||| |  
Db 10 LRLGFEVTL 19

## RESULT 14

A53357  
neurotoxin Bt-II - eastern Indian scorpion (fragment)  
C:Species: Mesobuthus tamulus (eastern Indian scorpion)  
C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 07-Feb-1997  
C:Accession: A53357  
R:Jalata, K.; Narayanan, P.  
Toxicon 32, 325-338, 1994  
A:Title: Purification, N-terminal sequence and structural characterization of a toxic protein  
A:Reference number: A53357; MUID:94267436  
A:Accession: A53357  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-24 <LAL>  
C:Keywords: neurotoxin

Query Match 17.7%; Score 26; DB 2; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TCRY 28  
| :||| |  
Db 18 TCRY 21

## RESULT 15

A32203  
dihydrofolate reductase (EC 1.5.1.3) - yeast (Candida albicans) (fragment)  
C:Species: Candida albicans  
C:Date: 20-Jul-1989 #sequence\_revision 20-Jul-1989 #text\_change 20-Mar-1996  
C:Accession: A32203  
R:Baccanari, D.P.; Tansik, R.L.; Joyner, S.S.; Fling, M.E.; Smith, P.L.; Freisheim, J.H.  
J. Biol. Chem. 264, 1100-1107, 1989  
A:Title: Characterization of Candida albicans dihydrofolate reductase.  
A:Reference number: A32203; MUID:89093095  
A:Accession: A32203  
A:Status: preliminary  
A:Molecule type: protein

A:Residues: 1-25 <BAC>  
C:Keywords: NADP; oxidoreductase

Query Match 17.7%; Score 26; DB 2; Length 25;  
Best Local Similarity 30.8%; Pred. No. 1.9e+03;  
Matches 4; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
QY 12 MAOMEKALSIGFE 24  
| :||| |  
Db 10 VVALPDLGLGK 22

Search completed: February 21, 2002, 17:12:55  
Job time: 111 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:12:40 ; Search time 10.02 Seconds  
(without alignments)  
102.457 Million cell updates/sec

Title: US-08-753-851-13  
Perfect score: 147  
Sequence: 1 LCRFNSLTPTMAQMEKALSIGFETCRY 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1617

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	19.7	24	1 PQQA_ACICA	P27532 acinetobact
2	29	19.7	24	1 PQQA_PSEFL	P55171 pseudomonas
3	27	18.4	28	1 APCL_RABIT	P33047 oryctolagus
4	26	17.7	16	1 RBL_CAPAN	P27063 capsicum an
5	26	17.7	16	1 RBL_CUCSA	P27064 cucumis sat
6	26	17.7	16	1 RBL_VIGSI	P27067 vigna sinen
7	26	17.7	24	1 SCX2_MESTA	P45868 mesobuthus
8	25	17.0	27	1 L52_ADE07	P05663 human adeno
9	24	16.3	22	1 VGLG_RABVA	P15199 rabies viru
10	24	16.3	25	1 Y194_ARCFU	O30045 archaeoglob
11	24	16.3	28	1 MAAL_RAT	P57113 rattus norv
12	23	15.6	25	1 PRLA_ACHLY	P27459 achromobact
13	23	15.6	26	1 RBL_VICFA	P05699 vicia faba
14	22.5	15.3	25	1 GBX1_MOUSE	P82976 mus musculu
15	22	15.0	10	1 TRP9_LEUMA	P81741 leucophaea
16	22	15.0	14	1 COCO_LIMPO	P35586 limulus pol
17	22	15.0	14	1 UC15_MAIZE	P80621 zea mays (m
18	22	15.0	15	1 KLOM_LUMTE	P11918 lumbricus t
19	22	15.0	25	1 LEOL_BIOGL	P80742 biophalarai
20	22	15.0	27	1 PETM_SKECO	O96810 skeletonema
21	22	15.0	27	1 DBH1_BIFLO	P17615 bifidobacte
22	22	15.0	28	1 VG9_SPV4	P11341 spiroplasma
23	21	14.3	10	1 TKL4_LOCMI	P30250 locusta mig
24	21	14.3	14	1 MAR1_ALTSP	P29399 alteromonas
25	21	14.3	20	1 YOAH_KLEAE	P56506 klebsiella
26	21	14.3	21	1 CFPA_TREPH	P56738 treponema p
27	21	14.3	21	1 PEDB_HYDAT	P80577 hydra atten
28	21	14.3	22	1 APCL_MACFA	P18657 macaca fasc
29	21	14.3	22	1 TX1_HETFU	P82850 heterometru
30	21	14.3	23	1 TX2_HETFU	P82851 heterometru
31	21	14.3	25	1 NP4_HUMAN	P18078 homo sapien
32	21	14.3	26	1 MEL_APIFL	P01504 apis florea
33	21	14.3	27	1 DCUP_RHOSH	P32920 rhodobacter

34 21 14.3 27 1 IAL7\_WHEAT P16852 triticum ae  
35 20 13.6 8 1 LCK1\_LEUMA P21140 leucophaea  
36 20 13.6 13 1 CRBL\_VESLE P17235 vespula lew  
37 20 13.6 20 1 CPXX\_RHORH P31718 rhodococcus  
38 20 13.6 21 1 MCT3\_MOUSE P21843 mus musculus  
39 20 13.6 23 1 PQQA\_KLEPN P27503 klebsiella  
40 20 13.6 24 1 ACHB\_ELEEL P09689 electrophor  
41 20 13.6 24 1 FIBG\_CANPA P12800 canis famil  
42 20 13.6 24 1 PSBR\_WHEAT P12358 triticum ae  
43 20 13.6 25 1 G3P2\_JACOR P80447 jaculus ori  
44 20 13.6 27 1 CXK7\_CONPU P56633 conus purpu  
45 20 13.6 27 1 DMS4\_PHYSA P80280 phyllomedus

ALIGNMENTS

RESULT 1  
PQQA\_ACICA STANDARD; PRT; 24 AA.  
AC P27532;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE COENZYME PQQ SYNTHESIS PROTEIN A (COENZYME PQQ SYNTHESIS PROTEIN IV).  
GN PQQA OR PQQIV.  
OS Acinetobacter calcoaceticus.  
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
OC Acinetobacter.  
OX NCBI\_TaxID=471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LMD 79.41;  
RX MEDLINE=89123056; PubMed=2536663;  
RA Goosen N., Horsman H.P.A., Huinen R.G.M., van de Putte P.;  
RT "Acinetobacter calcoaceticus genes involved in biosynthesis of the  
coenzyme pyrrolo-quinoline-quinone: nucleotide sequence and  
expression in Escherichia coli K-12.";  
RL J. Bacteriol. 171:447-455(1989).  
CC -I- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)  
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ  
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH  
CC ARE NECESSARY FOR THE SYNTHESIS.  
CC -I- SIMILARITY: TO OTHER BACTERIAL PQQA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X06452; -; NOT\_ANNOTATED\_CDS.  
DR PIR; B32252; B32252.  
KW PQQ.  
SQ SEQUENCE 24 AA; 2938 MW; 0E671EB9B7EA6838 CRC64;

Query Match 19.7%; Score 29; DB 1; Length 24;  
Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 19 LSGIFETCRY 28  
| | | | |  
Db 11 LRIGFEVTMY 20

RESULT 2  
PQQA\_PSEFL STANDARD; PRT; 24 AA.  
ID PQQA\_PSEFL  
AC P55171;  
DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE COENZYME PQQ SYNTHESIS PROTEIN A.  
GN PQQA.  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHA0;  
RX MEDLINE=96064397; PubMed=8526497;  
RA Schneider U., Keel C., Detago G., Haas D.;  
RT "Pns-directed cloning of pqq genes from Pseudomonas fluorescens CHA0:  
RT mutational inactivation of the genes results in overproduction of the  
RT antibiotic pyoluteorin.";  
RL Appl. Environ. Microbiol. 61:3856-3864(1995).  
CC -!- FUNCTION: REQUIRED FOR COENZYME PYROLO-QUINOLINE-QUINONE (PQQ)  
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ  
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH  
CC ARE NECESSARY FOR THE SYNTHESIS.  
CC -!- SIMILARITY: TO OTHER BACTERIAL PQQA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X87299; CAA60732.1;  
DR EMBL; X87299; CAA60731.1; ALT\_INIT.  
KW PQQ.  
SQ SEQUENCE 24 AA; 2871 MW; ED13A879B2EA3E8E CRC64;

Query Match 19.7%; Score 29; DB 1; Length 24;  
Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 19 LSIGFETCRY 28  
I I I I I I  
Db 11 LRIGFEVTMY 20

RESULT 3  
APCL\_RABIT  
ID APCL\_RABIT STANDARD; PRT; 28 AA.  
AC P33047;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE APOLIPROTEIN C-I (APO-CI) (FRAGMENT).  
GN APOC1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=91093092; PubMed=2266137;  
RA Weisgraber K.H., Mahley R.W., Kowal R.C., Herz J., Goldstein J.L.,  
RA Brown M.S.;  
RT "Apolipoprotein C-I modulates the interaction of apolipoprotein E  
RT with beta-migrating very low density lipoproteins (beta-VLDL) and  
RT inhibits binding of beta-VLDL to low density lipoprotein  
RT receptor-related protein.";  
RL J. Biol. Chem. 265:22453-22459(1990).  
CC -!- FUNCTION: APPEARS TO MODULATE THE INTERACTION OF APOE WITH BETA-  
CC MIGRATING VLDL AND INHIBIT BINDING OF BETA-VLDL TO THE LDL  
CC RECEPTOR-RELATED PROTEIN.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -!- TISSUE SPECIFICITY: SECRETED IN PLASMA.  
CC -!- SIMILARITY: BELONGS TO THE APOC1 FAMILY.  
DR PIR; A23691; A23691.  
KW HSP; P02654; 10PP.  
KW Plasma; Lipid transport; VLDL.  
FT NON\_TER 28 28  
SQ SEQUENCE 28 AA; 3165 MW; 2903C027F676541C CRC64;

Query Match 18.4%; Score 27; DB 1; Length 28;  
Best Local Similarity 46.2%; Pred. No. 7.1e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 KAFNSTLPTMAQM 15  
I I I I I I  
Db 15 KEFGNTLEAKRM 27

RESULT 4  
RBL\_CAPAN  
ID RBL\_CAPAN STANDARD; PRT; 16 AA.  
AC P27063;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
DE LARGE SUBUNIT) (FRAGMENT).  
GN RBL.  
OS Capsicum annuum (Bell pepper).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.  
OX NCBI\_TaxID=4072;  
RN [1]  
RP SEQUENCE, METHYLATION, AND ACETYLATION.  
RA Houtz R.L., Poneleit L., Jones S.B., Royer M., Stults J.T.;  
RT "Posttranslational modifications in the amino-terminal region of the  
RT large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from  
RT several plant species.";  
RL Plant Physiol. 98:1170-1174(1992).  
CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF  
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
CC ACTIVE SITE.  
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =  
CC 2 3-PHOSPHO-D-GLYCERATE.  
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =  
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.  
CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.  
DR InterPro: IPR000685; Rubisco\_large.  
DR PROSITE: PS00157; RUBISCO\_LARGE; PARTIAL.  
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;  
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation;  
KW Methylation.  
FT MOD\_RES 1 1 ACETYLATION.  
FT MOD\_RES 12 12 METHYLATION (TRF).  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1648 MW; D0AD959A82688872 CRC64;

Query Match 17.7%; Score 26; DB 1; Length 16;  
Best Local Similarity 45.5%; Pred. No. 5.7e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 14 QMEKALSIGFE 24  
I I I I I I  
Db 2 QTETKASVGFK 12



```

RESULT 5
RBL_CUCSA
ID RBL_CUCSA STANDARD; PRT; 16 AA.
AC P27064;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RUBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO
DE LARGE SUBUNIT) (FRAGMENT).
GN RBCL.
OS Cucumis sativus (Cucumber).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;
[1]
RN SEQUENCE, METHYLATION, AND ACETYLATION.
RA Houtz R.L., Poneleit L., Jones S.B., Royer M., Stults J.T.;
RT "Posttranslational modifications in the amino-terminal region of the
RT large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from
RT several plant species.";
RL Plant Physiol. 98:1170-1174(1992).
CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
CC 2 3-PHOSPHO-D-GLYCERATE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
DR InterPro: IPR000685; RUBISCO_large.
DR PROSITE: PS00157; RUBISCO_LARGE; PARTIAL.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation;
KW Methylation.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 12 12 METHYLATION (TRI-).
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1648 MW; D0AD959A82688872 CRC64;

Query Match 17.7%; Score 26; DB 1; Length 16;
Best Local Similarity 45.5%; Pred. No. 5.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 QMEKALSIGFE 24
DB 2 QTETKASVGFK 12

RESULT 6
RBL_VIGSI
ID RBL_VIGSI STANDARD; PRT; 16 AA.
AC P27067;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RUBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO
DE LARGE SUBUNIT) (FRAGMENT).
GN RBCL.
OS Vigna sinensis (Cowpea).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX NCBI_TaxID=3920;

```

```

[1]
RN SEQUENCE, METHYLATION, AND ACETYLATION.
RA Houtz R.L., Poneleit L., Jones S.B., Royer M., Stults J.T.;
RT "Posttranslational modifications in the amino-terminal region of the
RT large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from
RT several plant species.";
RL Plant Physiol. 98:1170-1174(1992).
CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
CC 2 3-PHOSPHO-D-GLYCERATE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
DR InterPro: IPR000685; RUBISCO_large.
DR PROSITE: PS00157; RUBISCO_LARGE; PARTIAL.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation;
KW Methylation.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 12 12 METHYLATION (TRI-).
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1648 MW; D0AD959A82688872 CRC64;

Query Match 17.7%; Score 26; DB 1; Length 16;
Best Local Similarity 45.5%; Pred. No. 5.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 QMEKALSIGFE 24
DB 2 QTETKASVGFK 12

RESULT 7
SCX2_MESTA
ID SCX2_MESTA STANDARD; PRT; 24 AA.
AC P45668;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROTOXIN II (BT-II) (FRAGMENT).
OS Mesobuthus tamulus (Eastern Indian scorpion) (Buthus tamulus).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34647;
[1]
RN SEQUENCE.
RP TISSUE=Venom;
RX MEDLINE=94287436; PubMed=8016854;
RA Lala K., Narayanan P.;
RT "Purification, N-terminal sequence and structural characterization of
RT a toxic protein from the Indian scorpion venom Buthus tamulus.";
RL Toxicon 32:325-338(1994).
CC -1- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION
CC OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
CC THIS TOXIN IS ACTIVE AGAINST MAMMALS. LD(50) IS 2.25 MG/KG IN MICE
CC BY SUBCUTANEOUS INJECTION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
KW Neurotoxin; Sodium channel inhibitor.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2686 MW; DA1990C8FF9E2769 CRC64;

Query Match 17.7%; Score 26; DB 1; Length 24;

```

```

Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TCYR 28
   IIII
Db 18 TCYR 21

RESULT 8
L52_ADE07
ID L52_ADE07 STANDARD; PRT; 27 AA.
AC P05663;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LATE L1 52 KDA PROTEIN (FRAGMENT).
OS Human adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GOMEN;
RX MEDLINE=83183660; PubMed=6301944;
RA Engler J.A., Hoppe M.S., van Bree M.P.;
RT "The nucleotide sequence of the genes encoded in early region 2b of
RL human adenovirus type 7.";
Gene 21:145-159(1983).
CC -!- FUNCTION: INVOLVED IN VIRION ASSEMBLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13357; CAA31737.1; -.
DR PIR; S07817; S07817.
KW Transmembrane; Envelope protein; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 >22 SPIKE GLYCOPROTEIN.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2469 MW; BBA59CE12F3FEF85 CRC64;

Query Match 16.3%; Score 24; DB 1; Length 22;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 19 LSIGFETC 26
   I : I I I
Db 10 LLLGFSLC 17

RESULT 10
Y194_ARCFU
ID Y194_ARCFU STANDARD; PRT; 25 AA.
AC Q30045;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN AF0194.
GN AF0194.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001093; AAB91054.1; -.
DR TIGR; AF0194; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 25 AA; 2828 MW; 5F480F03C367419F CRC64;

```

Query Match 16.3%; Score 24; DB 1; Length 25;

Best Local Similarity 58.3%; Pred. No. 1.8e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 15 MEKALSI--GFE 24

DB 11 LERAYRIEAGFE 22

# RESULT 11

MAAL\_RAT ID MAAL\_RAT STANDARD; PRT; 28 AA.  
 AC P57113;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MALEYLACETOATE ISOMERASE (EC 5.2.1.2) (MAAI) (GLUTATHIONE S-TRANSFERASE ZETA 1) (BC 2.5.1.18) (GSTz1-1) (FRAGMENT).  
 GN GSTZ1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 ON NCBI\_TaxID=101116;  
 RP SEQUENCE, AND CHARACTERIZATION.  
 RC STRAIN=FISCHER 344; TISSUE=Liver;  
 RX MEDLINE=98198370; PubMed=9531472;  
 RA Tong Z., Board P.G., Anders M.W.;  
 RT "Glutathione transferase zeta catalyzes the oxygenation of the  
 RT carcinogen dichloroacetic acid to glyoxylic acid.";  
 RL Biochem. J. 331:371-374(1998).  
 CC -1- FUNCTION: PROBABLE BIFUNCTIONAL ENZYME SHOWING MINIMAL  
 CC GLUTATHIONE-CONJUGATING ACTIVITY WITH ETHACRYNIC ACID AND 7-  
 CC CHLORO-4-NITROBENZ-2-OXA-1, 3-DIAZOLE AND MALEYLACETOATE  
 CC ISOMERASE ACTIVITY. HAS ALSO LOW GLUTATHIONE PEROXIDASE ACTIVITY  
 CC WITH T-BUTYL AND CUMENE HYDROPEROXIDES (BY SIMILARITY). IS ABLE TO  
 CC CATALYZE THE GLUTATHIONE DEPENDENT OXYGENATION OF DICHLOROACETIC  
 CC ACID TO GLYOXYLIC ACID.  
 CC -1- CATALYTIC ACTIVITY: 4-MALEYLACETOATE = 4-FUMARYLACETOATE.  
 CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-G.  
 CC -1- COFACTOR: THE MAAI ACTIVITY REQUIRES GLUTATHIONE (BY SIMILARITY).  
 CC -1- PATHWAY: CATABOLISM OF TYROSINE; FOURTH STEP, CATABOLISM OF  
 CC PHENYLALANINE; FIFTH STEP.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOSOLASMIC.  
 CC -1- PTM: THE N-TERMINUS IS BLOCKED (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ZETA FAMILY.  
 KW Isomerase; Transferase; Multifunctional enzyme;  
 KW Phenylalanine catabolism; Tyrosine catabolism.  
 FT NON\_TER 1 1  
 FT NON\_TER 28 28  
 SQ SEQUENCE 28 AA; 2943 MW; 1070608C44491C25 CRC64;

Query Match 16.3%; Score 24; DB 1; Length 28;

Best Local Similarity 50.0%; Pred. No. 2.1e+03;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 16 EKALSTGF 23

DB 2 QKAITSGF 9

# RESULT 12

PRLA\_ACHLY ID PRLA\_ACHLY STANDARD; PRT; 25 AA.  
 AC P27459;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ALPHA-LYTIC PROTEASE (EC 3.4.21.12) (ALPHA-LYTIC ENDOPEPTIDASE) (FRAGMENT).  
 OS Achromobacter lyticus.

Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;

Achromobacter.

NCBI\_TaxID=224;

RP SEQUENCE.

RC STRAIN=M497-1;

RX MEDLINE=91035265; PubMed=2228973;

RA Li S.L., Norioka S., Sakiyama F.;

"Molecular cloning and nucleotide sequence of the beta-lytic protease

gene from Achromobacter lyticus.";

J. Bacteriol. 172:6506-6511(1990).

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS, ESPECIALLY BONDS

ADJACENTS TO L-ALANINE RESIDUES.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A; ALSO KNOWN AS THE

ALPHA-LYTIC PROTEASE FAMILY.

DR HSP; P00778; IPI04.

DR MEROPS; S01.268; -.

DR InterPro: IPR001254; Trypsin.

DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.

DR PROSITE; PS00135; TRYPSIN\_SER; PARTIAL.

KW Hydrolase; Serine protease; Zymogen.

FT UNSURE 17 17

FT NON\_TER 25 25

SQ SEQUENCE 25 AA; 2572 MW; 33F56B45BF81573E CRC64;

## Query Match

Best Local Similarity 15.6%; Score 23; DB 1; Length 25;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 20 SIGFETCR 27

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

DB 18 SVGFSTVR 25

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X01167; CAA25614.1; -;  
 DR PIR; S07303; S07303.  
 DR HSSP; P00875; 1BUR.  
 DR Mendel; 5434; VICfa; rbcL;1.  
 DR InterPro; IPR000685; RUBISCO\_LARGE.  
 DR PROSITE; PS00157; RUBISCO\_LARGE; PARTIAL.  
 KW Photosynthesis; Carbon dioxide fixation; Photorespiration;  
 KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation.  
 FT PROPEP 1 2 BY SIMILARITY.  
 FT CHAIN 3 >26 RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE  
 FT CHAIN 3 3 ACETYLATION (BY SIMILARITY).  
 FT MOD\_RES 3 3  
 FT NON\_TER 26 26  
 SQ SEQUENCE 26 AA; 2955 MW; 754F486180E67E98 CRC64;

Query Match 15.68; Score 23; DB 1; Length 26;  
 Best Local Similarity 36.4%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 14 QMEKALSIGFE 24

DB 4 QTETKARVGQ 14

RESULT 14

ID GBX1\_MOUSE STANDARD; PRT; 25 AA.  
 AC P82976;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HOMEBOX PROTEIN GBX-1 (GASTRULATION AND BRAIN-SPECIFIC HOMEBOX  
 DE PROTEIN 1) (FRAGMENT).  
 GN GBX1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Spleen;  
 RX MEDLINE=92073357; PubMed=1720547;  
 RA Murtha M.T., Leckman J.F., Ruddle F.H.;  
 RT "Detection of homeobox genes in development and evolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715(1991).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M81662; AAA63311.1; -;  
 DR MGD; MGI:95667; Gbx1.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; PARTIAL.  
 DR PROSITE; PS50071; HOMEBOX\_2; PARTIAL.  
 KW Homeobox; DNA-binding; Nuclear protein; Transcription regulation.  
 FT DNA\_BIND 1 1  
 FT NON\_TER <1 >25 HOMEBOX.  
 FT NON\_TER 25 25  
 SQ SEQUENCE 25 AA; 2873 MW; FF273239ADF2F55B CRC64;

Query Match 15.3%; Score 22.5; DB 1; Length 25;  
 Best Local Similarity 35.0%; Pred. No. 3.1e+03;  
 Matches 7; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

OY 2 CKAFNSTLPTMAQMEKALSI 21  
 DB 1 CKKYL-SLTERSQIAHAKL 19

RESULT 15

ID TRP9\_LEUMA STANDARD; PRT; 10 AA.  
 AC P81741;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE TACHYKININ-RELATED PEPTIDE 9 (LEMRP 9).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberidae; Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Brain;  
 RX MEDLINE=97269266; PubMed=9114447;  
 RA Muren J.E., Naessel D.R.;  
 RT "Seven tachykinin-related peptides isolated from the brain of the  
 RT Madeira cockroach; evidence for tissue-specific expression of  
 RT isoforms".  
 RL Peptides 18:7-15(1997).  
 CC -1- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY  
 CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- MASS SPECTROMETRY: MW=1081.5; METHOD=MALDI.  
 CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
 KW Tachykinin; Neuropeptide; Amidation.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1081 MW; 9E469D66D9C87685 CRC64;

Query Match 15.0%; Score 22; DB 1; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 18 ALSIGFETCR 27

DB 1 APSMGFGQMR 10

Search completed: February 21, 2002, 17:15:40  
 Job time: 180 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:12:20 ; Search time 22.01 Seconds  
(without alignments)  
186,080 Million cell updates/sec

Title: US-08-753-851-13  
Perfect score: 147  
Sequence: 1 LKAFNSTLPTMAQMEKALSIGFETCRY 28

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 11317

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	87.8	25	4 Q9UC30.	Q9UC30 homo sapien
2	58	39.5	23	4 Q9UCB0	Q9UCB0 homo sapien
3	30.5	20.7	18	12 Q86380	Q86380 rubella vir
4	29	19.7	23	5 Q86705	Q86705 ilyanassa o
5	29	19.7	23	5 Q9BM01	Q9BM01 hydra litto
6	29	19.7	27	5 Q9Y192	Q9Y192 priapulid c
7	28	19.0	23	2 Q9F901	Q9F901 pseudomonas
8	28	19.0	27	5 Q26207	Q26207 phagocata w
9	28	19.0	27	5 Q27342	Q27342 stylaria la
10	28	19.0	27	5 Q9JUG3	Q9JUG3 steatoda tr
11	28	19.0	28	1 Q9UWM4	Q9UWM4 methanobact
12	27.5	18.7	19	10 Q9S882	Q9S882 chlamydomon
13	27	18.4	14	5 Q9NFK8	Q9NFK8 brigia paha
14	27	18.4	16	12 Q79455	Q79455 human immun
15	27	18.4	16	12 Q79461	Q79461 human immun
16	27	18.4	18	4 Q9UCL5	Q9UCL5 homo sapien
17	27	18.4	21	1 Q9UWH2	Q9UWH2 thermococcu
18	27	18.4	24	2 Q50436	Q50436 methylolaci
19	27	18.4	25	5 Q9BM19	Q9BM19 lissomyema

20	18.4	27	12	Q9WRL1	Q9WRL1 hepatitis b
21	17.7	15	8	Q95773	Q95773 conolophus
22	17.7	16	2	Q9K5A9	Q9K5A9 enterobacte
23	17.7	18	12	Q9QEX3	Q9QEX3 human immun
24	17.7	20	4	Q9UCE6	Q9UCE6 homo sapien
25	17.7	21	6	Q9TR36	Q9TR36 bos taurus
26	17.7	23	2	Q9ZAA0	Q9ZAA0 pseudomonas
27	17.7	23	4	Q9UCG9	Q9UCG9 homo sapien
28	17.7	23	5	Q25136	Q25136 halloctis ru
29	17.7	25	5	Q25873	Q25873 polycellul f
30	17.7	26	5	Q9NJR8	Q9NJR8 euprymna sc
31	17.7	26	5	Q9NJR5	Q9NJR5 euprymna sc
32	17.7	26	11	Q9OX07	Q9OX07 rattus norv
33	17.7	27	5	Q46193	Q46193 holopneuste
34	17.7	27	5	Q94450	Q94450 chaetopteru
35	17.7	27	5	Q26389	Q26389 ctenodrilus
36	17.7	27	5	Q23740	Q23740 ctenodrilus
37	17.7	27	5	Q06872	Q06872 branchiost
38	17.3	25	13	P82741	P82741 rana catesb
39	17.3	27	8	Q79901	Q79901 oplurus cuv
40	17.0	18	2	Q52411	Q52411 thermophil
41	17.0	20	4	Q9UC92	Q9UC92 homo sapien
42	17.0	20	4	Q9UC84	Q9UC84 homo sapien
43	17.0	22	13	Q9PS00	Q9PS00 micropogoni
44	17.0	23	5	Q9TWC4	Q9TWC4 dirofilaria
45	17.0	24	4	Q16397	Q16397 homo sapien

## ALIGNMENTS

## RESULT 1

Q9UC30 ID Q9UC30 PRELIMINARY; PRT; 25 AA.  
AC Q9UC30;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE 85 KDA GLIOMA MEMBRANE PROTEIN/CD44 HOMOLOG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95079448; PubMed=7527301;  
RA Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;  
RT "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes overexpressed CD44 in glioma cells.";  
RL Cancer Immunol. Immunother. 39:313-317(1994).  
DR InterPro: IPR000538; Link.  
DR ProDom: PD000918; Link: 1.  
SQ SEQUENCE 25 AA; 2810 MW; C65A2E285F780F4B CRC64;

Query Match 87.8%; Score 129; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY 4 AFNSTLPTMAQMEKALSIGFETCRY 28

|||||  
1 AFNSTLPTMAQMEKALSIGFETCRY 25

## RESULT 2

Q9UCB0 ID Q9UCB0 PRELIMINARY; PRT; 23 AA.  
AC Q9UCB0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE LYMPHOCYTE HOMING RECEPTOR CD44H (FRAGMENT).  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN NCBI\_TaxID=9606;  
RP [1]  
RX MEDLINE=94149816; PubMed=7508992;  
RA Shepley M.P., Racaniello V.R.;  
RT "A monoclonal antibody that blocks poliovirus attachment recognizes  
RL the lymphocyte homing receptor CD44.";  
RL J. Virol. 68:1301-1308(1994).  
SQ SEQUENCE 23 AA; 2573 MW; 2D6C726199AD402D CRC64;

Query Match 39.5%; Score 58; DB 4; Length 23;  
Best Local Similarity 92.3%; Pred. No. 0.025;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 EKALSIGFETCRY 28  
DB 1 EKALSIGFETCRY 13  
|||||  
|||||

RESULT 3  
Q86380 PRELIMINARY; PRT; 18 AA.  
AC Q86380;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE SUBGENOMIC INITIATION SITE RNA (FRAGMENT).  
OS Rubella virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;  
OC Rubivirus.  
OC NCBI\_TaxID=11041;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89085604; PubMed=2909990;  
RA Frey T.K., Marr L.D., Sanchez A., Simmons R.B.;  
RT "Identification of the 5' end of the rubella virus subgenomic RNA.";  
RL Virology 168:191-194(1989).  
DR EMBL; M23030; AAA72714.1; -.  
FT NON\_TER 18  
SQ SEQUENCE 18 AA; 1950 MW; 5D7CAF43BB769368 CRC64;

Query Match 20.7%; Score 30.5; DB 12; Length 18;  
Best Local Similarity 57.1%; Pred. No. 4.9e+02;  
Matches 8; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 7 STLP-TMAQMEKAL 19  
DB 3 STTPITMEDLOKAL 16  
|||||  
|||||

RESULT 4  
O96705 PRELIMINARY; PRT; 23 AA.  
AC O96705;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HOX1 PROTEIN (FRAGMENT).  
OS Ilyanassa obsoleta.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Muricoidae; Nassariidae; Ilyanassa.  
OX NCBI\_TaxID=34582;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Conrad A.H., Conrad G.W.;  
RT "Conserved transcription factor genes expressed in the veliger larvae  
RT of the gastropod mollusc, Ilyanassa obsoleta.";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF118831; AAD13788.1; -.

DR HSSP; P02833; 9ANT.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
FT NON\_TER 1  
FT NON\_TER 23  
SQ SEQUENCE 23 AA; 2665 MW; D0556DA25DA41458 CRC64;

Query Match 19.7%; Score 29; DB 5; Length 23;  
Best Local Similarity 38.9%; Pred. No. 1.1e+03;  
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 FNSTLPTMAQMEKALSIG 22  
DB 1 FNKYLTRARRIEIAASIG 18  
|||||  
|||||

RESULT 5  
Q9BM01 PRELIMINARY; PRT; 25 AA.  
AC Q9BM01;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE MARINER-LIKE TRANSPOSASE (FRAGMENT).  
OS Hydra littoralis.  
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;  
OC Hydridae; Hydra.  
OX NCBI\_TaxID=42240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TRANSPOSON=MARI.  
RX MEDLINE=20570504; PubMed=11121049;  
RA Arkhipova I., Meselson M.;  
RT "Transposable elements in sexual and asexual taxa.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).  
DR EMBL; AY014005; AAG59977.1; -.  
FT NON\_TER 1  
FT NON\_TER 25  
SQ SEQUENCE 25 AA; 2777 MW; CBF9DD69D7AD074C CRC64;

Query Match 19.7%; Score 29; DB 5; Length 25;  
Best Local Similarity 38.9%; Pred. No. 1.2e+03;  
Matches 7; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

QY 10 PTMAQ--MEKALSIGFET 25  
DB 3 PHVAQPTLRKLNALGYET 20  
|||||  
|||||

RESULT 6  
Q9Y192 PRELIMINARY; PRT; 27 AA.  
AC Q9Y192;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE LABIAL HOMEODOMAIN PROTEIN (FRAGMENT).  
GN LAB.  
OS Priapulus caudatus.  
OC Eukaryota; Metazoa; Priapulida; Priapulidae; Priapulidae.  
OX NCBI\_TaxID=37621;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93318125; PubMed=10391241;  
RA de Rosa R., Grenier J.G., Andreeva T., Cook C.E., Adoutte A., Akam M.;  
RT "Hox genes in brachiopods and priapulids and protostome evolution.";  
RL Nature 399:772-776(1999).  
DR EMBL; AF144884; AAD40640.1; -.  
DR HSSP; P02833; 9ANT.  
DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 FT NON\_TER 27  
 SQ SEQUENCE 27 AA; 3143 MW; 888ED479D6DA25D CRC64;

Query Match 19.7%; Score 29; DB 5; Length 27;  
 Best Local Similarity 38.9%; Pred. No. 1.3e+03;  
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 FNSTLPTMAQMEKALSIG 22  
 || | :| | | :|  
 Db 2 FNKYLTRARRIEIAASLG 19

RESULT 7  
 Q9F9U1 PRELIMINARY; PRT; 23 AA.  
 AC Q9F9U1  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE COENZYME PQQ SYNTHESIS PROTEIN A.  
 GN PQQA.  
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=316;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BC-2;  
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,  
 RA Terwilliger T.C.;  
 RT "Identification and characterization of genes activated by 2-  
 RT chloroethanol in Pseudomonas stutzeri BC-2.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF176640; AAG09250.1;  
 SQ SEQUENCE 23 AA; 2809 MW; 80815999B7EBCC47 CRC64;

Query Match 19.0%; Score 28; DB 2; Length 23;  
 Best Local Similarity 60.0%; Pred. No. 1.6e+03;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 19 LSTGFETCRY 28  
 | |||| |  
 Db 10 LRIGFEVTLY 19

RESULT 8  
 Q26207 PRELIMINARY; PRT; 27 AA.  
 AC Q26207  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE HOMEBOX PROTEIN (FRAGMENT).  
 OS Phagocata woodworthi.  
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Turbellaria;  
 OC Tricladida; Paludicola; Planariidae; Phagocata.  
 OX NCBI\_TaxID=27900;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94297796; PubMed=7912986;  
 RA Bartels J.L., Murtha M.T., Ruddle F.H.;  
 RT "Multiple Hox/HOM-class homeoboxes in Platyhelminthes.";  
 RL Mol. Phylogenet. Evol. 2:143-151(1993).  
 DR EMBL; LI9217; AAB39344.1;  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.

KW Homeobox; DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 FT NON\_TER 27  
 SQ SEQUENCE 27 AA; 3164 MW; 293591649C6DB3BD CRC64;

Query Match 19.0%; Score 28; DB 5; Length 27;  
 Best Local Similarity 38.9%; Pred. No. 1.9e+03;  
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 FNSTLPTMAQMEKALSIG 22  
 || | :| | | :|  
 Db 2 FNQYLTRARRIEIATSMG 19

RESULT 9  
 Q27342 PRELIMINARY; PRT; 27 AA.  
 AC Q27342  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE LABIAL-TYPE GENE LAB2 HOMOLOG (FRAGMENT).  
 OS Stylaria lacustris.  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
 OC Tubificina; Naididae; Stylaria.  
 OX NCBI\_TaxID=41694;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95211276; PubMed=7697193;  
 RA Snow P., Buss L.W.;  
 RT "Hox/Hox-type homeoboxes from Stylaria lacustris (Annelida;  
 RT Oligochaeta).";  
 RL Mol. Phylogenet. Evol. 3:360-364(1994).  
 DR EMBL; S76354; AAB33895.1;  
 DR EMBL; S76352; AAB73975.1;  
 DR HSSP; P02833; 9ANT.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 KW Homeobox; Nuclear protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 27 AA; 3171 MW; 7907D4749D6DA24F CRC64;

Query Match 19.0%; Score 28; DB 5; Length 27;  
 Best Local Similarity 38.9%; Pred. No. 1.9e+03;  
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 FNSTLPTMAQMEKALSIG 22  
 || | :| | | :|  
 Db 2 FNRYLTRARRIEIAASLG 19

RESULT 10  
 Q9NJG3 PRELIMINARY; PRT; 27 AA.  
 AC Q9NJG3  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE LABIAL (FRAGMENT).  
 GN LAB.  
 OS Steatoda triangulosa.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Araneae;  
 OC Araneomorphae; Entelegynae; Araneioidea; Therididae; Steatoda.  
 OX NCBI\_TaxID=114396;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Abzhanov A., Popadic A., Kaufman T.C.;  
 RT "Chelicerate Hox genes and the homology of arthropod segments.";  
 RL Evol. Dev. 0:0-0(2000).  
 DR EMBL; AF151987; AAF73203.1;



DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 FT NON\_TER 1  
 FT NON\_TER 27  
 SQ SEQUENCE 27 AA; 3159 MW; 888ED462476DA25D CRC64;

Query Match 19.0%; Score 28; DB 5; Length 27;  
 Best Local Similarity 38.9%; Pred. No. 1.9e+03;  
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 FNSTLPTMAQMEKALSLG 22  
 || | :|| | | |  
 Db 2 FNKYLTRARIEIASLGL 19

RESULT 11  
 Q9UWM4  
 ID Q9UWM4 PRELIMINARY; PRT; 28 AA.  
 AC Q9UWM4;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE POLYFERREDOXIN-MVHB GENE PRODUCT.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145262;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92183831; PubMed=1312016;  
 RA Hedderich R., Albracht S.P., Koch J., Thauer R.K.;  
 RT "Isolation and characterization of polyferredoxin from  
 RT Methanobacterium thermoautotrophicum. The mvhb gene product of the  
 RT methylovigen-reducing hydrogenase operon.";  
 RL FEBS Lett. 298:65-68(1992).  
 DR HSSP: P55907; 1XER.  
 DR InterPro: IPR001450; 4FE4S\_ferredoxin.  
 DR Pfam: PF00037; fer4; 1.  
 DR PROSITE: PS00198; 4FE4S\_FERREDOXIN; 1.  
 SQ SEQUENCE 28 AA; 2923 MW; D22257BF296D235 CRC64;

Query Match 19.0%; Score 28; DB 1; Length 28;  
 Best Local Similarity 50.0%; Pred. No. 2e+03;  
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CKAFNSTLPTMA 13  
 | | | | | |  
 Db 12 CGACQGTCTAA 23

RESULT 12  
 Q9S882  
 ID Q9S882 PRELIMINARY; PRT; 19 AA.  
 AC Q9S882;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
 DE ATP SYNTHASE SUBUNIT II-B' (FRAGMENT).  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96128220; PubMed=8543042;  
 RA Fiedler H.R., Schmid R., Leu S., Shavit N., Strotmann H.;  
 RT "Isolation of CF0CF1 from Chlamydomonas reinhardtii cw15 and the N-  
 RT terminal amino acid sequences of the CF0CF1 subunits.";  
 RL FEBS Lett. 377:163-166(1995).  
 SQ SEQUENCE 19 AA; 2081 MW; A0AC64A247D406A2 CRC64;

Query Match 18.7%; Score 27.5; DB 10; Length 19;  
 Best Local Similarity 53.3%; Pred. No. 1.6e+03;  
 Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 3 KAFNSTLPTMAQME 16  
 | | :|| | | |  
 Db 4 KIFDNTLPVMAGEE 18

RESULT 13  
 Q9NFK8  
 ID Q9NFK8 PRELIMINARY; PRT; 14 AA.  
 AC Q9NFK8;  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE MMCI PROTEIN (FRAGMENT).  
 GN MMCI.  
 OS Brugia pahangi.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Brugia.  
 OX NCBI\_TaxID=6280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ames R.D., Thompson F., Devaney E.;  
 RT "A novel mRNA up-regulated in mammalian-derived microfilaria of  
 RT Brugia.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ277990; CAB93515.1;  
 FT NON\_TER 1  
 FT NON\_TER 14  
 SQ SEQUENCE 14 AA; 1602 MW; 5C4C62C55AB89397 CRC64;

Query Match 18.4%; Score 27; DB 5; Length 14;  
 Best Local Similarity 42.9%; Pred. No. 1.4e+03;  
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 AFNSTLPTMAQMEK 17  
 | | :|| | | |  
 Db 1 AFSATSLMMKQ 14

RESULT 14  
 Q79455  
 ID Q79455 PRELIMINARY; PRT; 16 AA.  
 AC Q79455;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE LIGHT-CHAIN COMPLEMENTARITY-DETERMINING REGION 3 MRNA (CLONE 14),  
 DE PARTIAL CDS (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92052225; PubMed=1719545;  
 RA Burton D.R., Barbas C.F. III, Persson M.A., Koenig S., Chanock R.M.,  
 RA Lerner R.A.;  
 RT "A large array of human monoclonal antibodies to type 1 human  
 RT immunodeficiency virus from combinatorial libraries of asymptomatic  
 RT seropositive individuals.";  
 RL proc. Natl. Acad. Sci. U.S.A. 88:10134-10137(1991).  
 DR EMBL; M80720; AAA44780.1;  
 DR HSSP; P01607; IREI.  
 FT NON\_TER 1  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1883 MW; 9015605C32B3D65E CRC64;

Query Match 18.4%; Score 27; DB 12; Length 16;

Best Local Similarity 38.5%; Pred. No. 1.6e+03;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 CKAFNSTLPTMAQ 14  
I: :||| |  
Db 3 CQKYSAPRTFGQ 15

RESULT 15

Q79461  
ID Q79461 PRELIMINARY; PRT; 16 AA.  
AC Q79461;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE LIGHT-CHAIN COMPLEMENTARITY-DETERMINING REGION 3 MRNA (CLONE 31),  
DE PARTIAL CDS (FRAGMENT).  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92052225; PubMed=1719545;  
RA Burton D.R., Barbas C.F. III, Persson M.A., Koenig S., Chanock R.M.,  
RA Lerner R.A.;  
RT "A large array of human monoclonal antibodies to type 1 human  
RT immunodeficiency virus from combinatorial libraries of asymptomatic  
RT seropositive individuals."  
RL Proc. Natl. Acad. Sci. U.S.A. 88:10134-10137(1991).  
DR EMBL; M80726; AAA44786.1; -;  
FT NON\_TER 1  
FT NON\_TER 16  
SQ SEQUENCE 16 AA; 1883 MW; 9015605C32B3D65E CRC64;

Query Match 18.4%; Score 27; DB 12; Length 16;  
Best Local Similarity 38.5%; Pred. No. 1.6e+03;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 CKAFNSTLPTMAQ 14  
I: :||| |  
Db 3 CQKYSAPRTFGQ 15

Search completed: February 21, 2002, 17:15:24  
Job time: 184 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:09:04 ; Search time 23.75 Seconds  
(without alignments)  
87.329 Million cell updates/sec

Title: US-08-753-851-13  
Perfect score: 147  
Sequence: 1 LCKAFNSTLPTMAQKEKALSIGFETCRY 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 231188

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	28	15 AAR53483	CD44 peptide CD44-
2	42	28.6	25	11 AAR06539	Human B-cell simul
3	40	27.2	18	16 AAR77857	Mycoplasma pirum a
4	38	25.9	22	16 AAR72687	Peptide #4 having
5	33.5	22.8	23	9 AAP80569	Peptide region of
6	33	22.4	15	20 AAW73686	M. tuberculosis an
7	33	22.4	15	20 AAW73796	M. tuberculosis an
8	33	22.4	24	20 AAY36359	Fragment of human
9	33	22.4	28	14 AAR34568	Domain 1 from inte
10	32.5	22.1	20	21 AAB13658	C. pneumoniae Swib
11	32.5	22.1	20	22 AAG83226	Chlamydia Swib pep

12	32	21.8	20	22	AAE03101	Human gene 13 enco
13	31.5	21.4	22	21	AAB13659	C. pneumoniae Swib
14	31.5	21.4	22	22	AAG83227	Chlamydia Swib pep
15	31	21.1	18	20	AAV19410	Lecithin:cholester
16	31	21.1	18	20	AAV19156	Lecithin:cholester
17	31	21.1	18	20	AAV18902	Lecithin:cholester
18	31	21.1	18	20	AAV18639	Lecithin:cholester
19	31	21.1	26	20	AAV45348	Human secreted pro
20	30.5	20.7	20	20	AAV29363	Original 3D profil
21	30.5	20.7	21	12	AAR11136	Rubella virus Ther
22	30	20.4	10	21	AAV85439	Tumour antigen pro
23	30	20.4	15	11	AAAR03455	Accessory moiety,
24	30	20.4	15	20	AAW73685	M. tuberculosis an
25	30	20.4	15	20	AAW73795	M. tuberculosis an
26	30	20.4	16	18	AAW25441	Grb2 N-terminal SH
27	30	20.4	16	22	AAB46168	Plasmodium sp mala
28	30	20.4	16	22	AAB49067	Malaria CS T3 T-ce
29	30	20.4	17	16	AAV78919	Malaria circumspor
30	30	20.4	17	16	AAV70911	Malaria circumspor
31	30	20.4	17	21	AAV99032	HLA class II bindi
32	30	20.4	18	21	AAB34521	Human secreted pro
33	30	20.4	18	21	AAV49259	CD4+ T cell epitop
34	30	20.4	19	19	AAW74524	Human brain zonula
35	30	20.4	19	21	AAV99033	HLA class II bindi
36	30	20.4	20	15	AAV58967	Cancer treating, a
37	30	20.4	21	10	AAV91504	Sequence of modifi
38	30	20.4	21	14	AAV45053	Basic (positively
39	30	20.4	21	14	AAV36312	Amphiphilic ion ch
40	30	20.4	21	14	AAV31083	C-terminal substd.
41	30	20.4	21	14	AAV35301	Amphiphilic peptid
42	30	20.4	21	14	AAV39015	Biologically activ
43	30	20.4	21	15	AAV55907	Ion channel formin
44	30	20.4	21	15	AAV50483	Amphiphilic pepti
45	30	20.4	21	15	AAV50367	Amphiphilic peptid

#### ALIGNMENTS

RESULT 1  
AAR53483  
ID AAR53483 standard; peptide: 28 AA.  
XX  
AC AAR53483;  
XX  
DT 01-DEC-1994 (first entry)  
XX  
CD44 peptide CD44-10.  
XX  
Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autolimmune disease;  
KW immunosuppressive.  
XX  
Homo sapiens.  
XX  
PN W09409811-A.  
XX  
11-MAY-1994.  
XX  
29-OCT-1993; 93WO-US10412.  
XX  
30-OCT-1992; 92US-0973339.  
PR  
(UYDU-) UNIV DUKE.  
XX  
Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
XX  
WPI; 1994-167121/20.  
DR  
Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis  
XX

PS Claim 4; Page 14; 83pp; English.

CC The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 52-79 of the CD44 protein sequence.

XX Sequence 28 AA;

Query Match 100.0%; Score 147; DB 15; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1.9e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKAFNSTLPTMAQMEKALSIGFETCRY 28

DB 1 lckafnstlptmaqmekalsigfctcry 28

RESULT 2

AAR06539

ID AAR06539 standard; protein; 25 AA.

XX AC AAR06539;

XX DT 04-JAN-1991 (first entry)

XX DE Human B-cell stimulator 2 (BSF2) antagonist.

XX KW Immunodeficiency; myeloma; chronic articular rheumatism;  
XX KW endotoxin shock;

XX OS Homo sapiens.

XX PN JP02188600-A.

XX PD 24-JUL-1990.

XX PF 17-JAN-1989; 89JP-0007944.

XX PR 17-JAN-1989; 89JP-0007944.

XX PA (CHUS ) CHUGAI PHARMACEUTICAL KK.

XX DR WPI; 1990-266224/35.

XX PT Polypeptide with antagonist against human B cell stimulator 2 -  
XX PT comprises human B cell stimulator with deficiency of 2 or more  
XX PT aminoacid(s) at N-and/or C-terminal.

XX PS Claim 12; Page 968; 6pp; Japanese.

XX CC BSF2 antagonists are useful in diagnosis and treatment of immuno-deficiencies eg. myeloma, chronic articular rheumatism and endotoxin shock. Antagonist comprises BSF2 peptide (pref. AAs 20-40) with substitutions of two or more C- and/or N-terminal AAs.

XX SQ Sequence 25 AA;

Query Match 28.6%; Score 42; DB 11; Length 25;  
Best Local Similarity 43.5%; Pred. No. 7.9;  
Matches 10; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 6 NSTLPTMAQMEKALSIGF--ETC 26

DB 2 nlnlptmaekdgcfqsgfneetc 24

RESULT

AAR77857

ID AAR77857 standard; Protein; 18 AA.

XX AC AAR77857;

XX DT 08-SEP-1995 (first entry)

XX DE Mycoplasma pirum adhesin immunogenic peptide (1055-1072).

XX KW Adhesin; cytoadherence; Mycoplasma; human immunodeficiency virus;  
XX KW HIV; vaccine; immunogen; antigen; infection.

XX OS Mycoplasma pirum.

XX PN FR2705970-A.

XX PD 09-DEC-1994.

XX PF 04-JUN-1993; 93FR-0006745.

XX PR 04-JUN-1993; 93FR-0006745.

XX PA (INSP ) INST PASTEUR.

XX PI Bahraoui E, Blanchard A, Ferris S, Montagnier L;  
XX PI Tham TN;

XX DR WPI; 1995-024735/04.

XX PT Nucleic acid encoding Mycoplasma pirum adhesin - and related  
XX PT proteins and antibodies, useful in diagnosis, treatment and  
XX PT prevention of M.pirum and HIV infection

XX PS Claim 7; Page 23; 55pp; French.

XX CC The M.pirum adhesin gene was isolated using a probe based on the 3'-  
XX CC end of the known M.genitalium adhesin gene. M.pirum is found in  
XX CC AIDS patients and the adhesin may have a role in infection of cells  
XX CC by HIV. If so, the M.pirum adhesin protein will be useful in vaccines  
XX CC to protect against the cytopathic effect of the Mycoplasma and against  
XX CC that of HIV. Specifically, antibodies raised against the peptides  
XX CC R77856 and R77857 (amino acids 201-219 and 1055-1072 of the adhesin,  
XX CC respectively) can inhibit infection of human cells by M.pirum and  
XX CC infection of human T lymphocytes by HIV.

XX SQ Sequence 18 AA;

Query Match 27.2%; Score 40; DB 16; Length 18;  
Best Local Similarity 53.8%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 MAQMEKALSIGFPE 24

DB 5 makhkkaikvgfe 17

RESULT 4

AAR72687

ID AAR72687 standard; peptide; 22 AA.

XX AC AAR72687;

XX DT 19-DEC-1995 (first entry)

XX DE Peptide #4 having inflammation site affinity.

XX KW Affinity; inflammation; radioactive label; imaging; diagnosis.

XX OS Synthetic.

XX PN EP649857-A1.

XX PD 26-APR-1995.

XX 20-OCT-1994; 94EP-0116583.  
 XX 22-OCT-1993; 93JP-0287752.  
 XX (NIMD) NIHON MEDIPHYSICS CO LTD.  
 PA Hanaoka K, Itaya Y, Shirakami Y;  
 XX WPI; 1995-156718/21.

XX New peptide(s) with affinity for inflammation sites - and  
 PT radio:labelled derivs. useful as diagnostic agents

XX Claim 1; Page 11; 21pp; English.

XX The sequences given in AAR72684-89 are peptides which have an affinity  
 CC to sites of inflammation. These peptides may be radioactively  
 CC labelled and used for imaging inflammation sites as diagnostic  
 CC agents. Tests on rats showed that these peptides would be safe at  
 CC levels upto 1000 times the recommended clinical dose. The peptides  
 CC were found to accumulate at sites of inflammation and remain there  
 CC long enough to allow effective imaging.

XX Sequence 22 AA;

Query Match 25.9%; Score 38; DB 16; Length 22;  
 Best Local Similarity 47.4%; Pred. No. 29;  
 Matches 9; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 2 CKAFNSTLPTMAQMEKALS 20  
 DB || | || | : || : |  
 5 ckvsnkalp--apiektis 21

RESULT 5  
 ID AAP80569 standard; protein; 23 AA.

XX AAP80569;

XX 07-NOV-1990 (first entry)

XX Peptide region of human immunodeficiency virus-1 GAG 3.

XX HIV-1; peptide; HIV-1 GAG 3; antibody; immunogen.

XX EP284587-A.

XX 28-SEP-1988.

XX 28-MAR-1988; 88EP-0850105.

XX 17-MAR-1987; 87SE-0701294.

XX 18-MAY-1987; 87US-0051726.

XX (VIRO-) VIROVAHL SA.

XX Vahine A, Svennerholm B, Rymo L, Jeansson S, Horal P;

XX WPI; 1988-272997/39.

XX Synthetic peptide antigens for detection of HIV-1 infection -  
 PT also useful as immunogens in vaccine compsns.

XX Claim 8; Page 25; 31pp; English.

XX The synthetic peptide correspond to regions of immunologically important  
 CC proteins of HIV-1. The peptide provides a superior, sensitive and  
 CC selective assay for the presence of antibodies to HIV-1. It may also be  
 CC used as an immunogen to elicit the prodn. of anti-HIV-1 antibodies.  
 CC X is the H of the N-terminal NH2 group or an additional amino acid

CC selected to facilitate coupling of the peptide to a carrier protein.  
 CC Y is absent or Cys, and Z is OH or NH2.  
 XX  
 SQ Sequence 23 AA;

Query Match 22.8%; Score 33.5; DB 9; Length 23;  
 Best Local Similarity 40.9%; Pred. No. 1.6e+02;  
 Matches 9; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 3 KAFNSTLPTMAQMEKALSIGFE 24  
 DB ||| : | : | ||| : :  
 5 kafs---pevipmtfsalsegvyz 23

RESULT 6

AAW73686  
 ID AAW73686 standard; Peptide; 15 AA.

XX AAW73686;

XX 24-MAR-1999 (first entry)

XX M. tuberculosis antigen clone ORF-1 peptide 1-10.

XX Antigen; M.tuberculosis protein; immunotherapy; tuberculosis; diagnosis;  
 KW infection.

XX Mycobacterium tuberculosis.

XX WO9853076-A2.

XX 26-NOV-1998.

XX 20-MAY-1998; 98WO-US10514.

XX 05-MAY-1998; 98US-0073009.

XX 20-MAY-1997; 97US-0858998.

XX (CORI-) CORIXA CORP.

XX Alderson MR, Campos-Neto A, Dillon DC, Skeiky YAW;

XX WPI; 1999-045315/04.

XX New isolated Mycobacterium tuberculosis antigens - used to develop  
 PT products for the prevention, treatment and diagnosis of tuberculosis  
 PT infection

XX Example 1; Page 53; 104pp; English.

XX This sequence represents an immunogenic portion of a Mycobacterium  
 CC tuberculosis antigen of the invention. The polypeptides are useful for  
 CC immunotherapy to treat or prevent tuberculosis (especially in humans);  
 CC e.g. they can be included with an acceptable carrier in pharmaceutical  
 CC compositions or included in vaccines, and administered to induce  
 CC protective immunity in a patient against M. tuberculosis. Tuberculosis is  
 CC a chronic, infectious disease generally caused by M. tuberculosis  
 CC infection, and if left untreated typically results in serious  
 CC complications and death. Fusion proteins containing the antigen, or DNA  
 CC molecules can similarly be included with an acceptable carrier in  
 CC pharmaceutical compositions or in vaccines and administered as above. The  
 CC polypeptides are also useful for diagnosis of tuberculosis, by contacting  
 CC dermal cells with at least one polypeptide and detecting an immune  
 CC response (especially induration) on the patient's skin. Inhibiting the  
 CC spread of tuberculosis requires vaccination and accurate diagnosis, since  
 CC antibiotic therapy may not be effective due to the existence of an  
 CC asymptomatic but contagious stage and to patient non-compliance. The  
 CC polypeptides overcome concerns of safety and efficacy of current  
 CC vaccination with live bacteria (usually Bacillus Calmette-Guerin) and  
 CC lack of sensitivity and specificity of existing diagnostic techniques.  
 XX Sequence 15 AA;

Query Match 22.4%; Score 33; DB 20; Length 15;  
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 TLPTMAQMEKA 18  
 Db 1 sldtmaqmqnqa 11

RESULT 7  
 AAW73796  
 ID AAW73796 standard; Peptide; 15 AA.

AC AAW73796;  
 XX  
 DT 24-MAR-1999 (first entry)

DE M. tuberculosis antigen clone ORF-1 peptide 1-10.

KW Antigen; M. tuberculosis protein; immunotherapy; tuberculosis; diagnosis;  
 KW infection.

OS Mycobacterium tuberculosis.

XX WO9853075-A2.

XX 26-NOV-1998.

PF 20-MAY-1998; 98WO-US10407.

PR 05-MAY-1998; 98US-0073010.

PR 20-MAY-1997; 97US-0859381.

PA (CORI-) CORIXA CORP.

PI Alderson MR, Campos-Neto A, Dillon DC, Skeiky YAW;

DR WPI; 1999-045314/04.

XX Polypeptide comprising immunogenic Mycobacterium tuberculosis  
 PT antigen - useful for immunisation against M. tuberculosis infection  
 PT to treat or prevent tuberculosis, and in diagnosis of tuberculosis

PS Example 1; Page 52; 100pp; English.

CC This sequence represents an immunogenic portion of a Mycobacterium  
 CC tuberculosis antigen of the invention. The polypeptides are useful for  
 CC immunotherapy to treat or prevent tuberculosis (especially in humans),  
 CC e.g., they can be included with an acceptable carrier in pharmaceutical  
 CC compositions or included in vaccines, and administered to induce  
 CC protective immunity in a patient against M. tuberculosis. Tuberculosis is  
 CC a chronic, infectious disease generally caused by M. tuberculosis  
 CC infection, and if left untreated typically results in serious  
 CC complications and death. Fusion proteins containing the antigen, or DNA  
 CC molecules can similarly be included with an acceptable carrier in  
 CC pharmaceutical compositions or in vaccines and administered as above. The  
 CC polypeptides are also useful for diagnosis of tuberculosis, by contacting  
 CC dermal cells with at least one polypeptide and detecting an immune  
 CC response (especially induration) on the patient's skin. Inhibiting the  
 CC spread of tuberculosis requires vaccination and accurate diagnosis, since  
 CC antibiotic therapy may not be effective due to the existence of an  
 CC asymptomatic but contagious stage and to patient non-compliance. The  
 CC polypeptides overcome concerns of safety and efficacy of current  
 CC vaccination with live bacteria (usually Bacillus Calmette-Guerin) and  
 CC lack of sensitivity and specificity of existing diagnostic techniques.

XX Sequence 15 AA;

Query Match 22.4%; Score 33; DB 20; Length 15;  
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 8 TLPTMAQMEKA 18  
 Db 1 sldtmaqmqnqa 11

RESULT 8  
 AAY36359  
 ID AAY36359 standard; Protein; 24 AA.

XX AAY36359;

XX 17-SEP-1999 (first entry)

DE Fragment of human secreted protein encoded by gene 3.

KW Human; secreted protein; cancer; tumour; developmental abnormality;  
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;  
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;  
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;  
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;  
 KW digestive disorder; endocrine disorder; infection; AIDS.

XX Homo sapiens.

XX WO9931117-A1.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-US27059.

PR 19-DEC-1997; 97US-0068369.

PR 18-DEC-1997; 97US-0068006.

PR 18-DEC-1997; 97US-0068008.

PR 18-DEC-1997; 97US-0068053.

PR 18-DEC-1997; 97US-0068054.

PR 18-DEC-1997; 97US-0068057.

PR 18-DEC-1997; 97US-0068064.

PR 18-DEC-1997; 97US-0070923.

PR 19-DEC-1997; 97US-0068169.

PR 19-DEC-1997; 97US-0068365.

PR 19-DEC-1997; 97US-0068367.

PR 19-DEC-1997; 97US-0068368.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;

PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;

PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;

XX Yu G;

DR WPI; 1999-418749/35.

XX New isolated human genes encoding secreted polypeptides

XX Disclosure; Page 420; 537pp; English.

XX AAX97916 to AAX98029 represent 110 isolated human secreted protein  
 CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by  
 CC the 110 human genes. The genes and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new genes.  
 CC Specific uses are described for each of the 110 genes, based on which  
 CC tissues they are most highly expressed in, and include developing  
 CC products for the diagnosis or treatment of cancer, tumours, developmental  
 CC abnormalities and foetal deficiencies, blood disorders, diseases of the  
 CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's  
 CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,  
 CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular

CC disorders, kidney disorders, digestive/endocrine disorders, infections  
 CC and AIDS. The polypeptides are also useful for identifying their binding  
 CC partners. The sequences given in RAX97907 to AAX97915 and AAY36223 are  
 CC used in the exemplification of the present invention.

XX SQ Sequence 24 AA;

Query Match 22.4%; Score 33; DB 20; Length 24;

Best Local Similarity 60.0%; Pred. No. 2e+02; Mismatches 2; Indels 0; Gaps 0;

QY 10 PTMAQMEKAL 19  
 I: |||||:

Db 15 psraqmekcm 24

RESULT 9

AAR34568

ID AAR34568 standard; Protein; 28 AA.

XX AC AAR34568;

XX 04-JUN-1993 (first entry)

XX Domain 1 from integrase like protein cre.

XX cytotoxic T-lymphocyte response; transformed Mycobacteria; BCG;  
 KW Mycobacterium smegmatis; vaccine; cell mediated immunity; HIV;  
 KW pertussis; malaria; influenza virus; CTL; herpes virus.  
 XX WO9307897-A.

XX 29-APR-1993.

XX 21-OCT-1992; 92WO-US09075.

XX 21-OCT-1991; 91US-0780261.

XX (MEDI-) MEDIMMUNE INC.

XX Stover CK;

XX WPI; 1993-152187/18.

XX Expression vector for expressing protein or polypeptide in  
 PT mycobacterium - contg DNA sequences encoding lipoprotein  
 PT secretion signal and peptide heterologous to bacteria expressing  
 PT fusion protein of lipoprotein heterologous to bacteria  
 XX Disclosure; Fig 29; 86pp; English.

XX This sequence is domain 1 from the integrase like protein cre. The  
 CC whole coding sequence of the enzyme could be included in an expression  
 CC vector in order to aid its integration into its host. The  
 CC expression vector also includes at least the secretion signal of a  
 CC lipoprotein and a second sequence encoding a heterologous protein  
 CC and a mycobacterial promoter to control expression of the  
 CC heterologous protein.

XX SQ Sequence 28 AA;

Query Match 22.4%; Score 33; DB 14; Length 28;

Best Local Similarity 70.0%; Pred. No. 2.4e+02; Mismatches 2; Indels 0; Gaps 0;

QY 13 AQMEKALSIG 22  
 I: |||||:

Db 2 agvekalsig 11

RESULT 10

AAB13658

ID AAB13658 standard; Peptide; 20 AA.

XX AC AAB13658;

XX 02-FEB-2001 (first entry)

XX C. pneumoniae Swib peptide 61-80.

XX Chlamydial infection; sexually transmitted disease;

KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
 KW trachoma; blindness; acute respiratory tract infection;  
 KW atherosclerosis; coronary heart disease; antibacterial.

XX Chlamydia pneumoniae.

XX WO200034483-A2.

XX 15-JUN-2000.

XX 08-DEC-1999; 99WO-US29012.

XX 08-DEC-1998; 98US-0208277.

XX 08-APR-1999; 99US-0288594.

XX 01-OCT-1999; 99US-0410568.

XX 22-OCT-1999; 99US-0426571.

XX (CORI-) CORIXA CORP.

XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;

XX WPI; 2000-431303/37.

XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
 PT comprises immunogenic portion of Chlamydia antigen, which comprises  
 PT amino acid sequence encoded by polynucleotide sequence -

XX Disclosure; Page 229; 256pp; English.

XX The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydiae are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.

XX SQ Sequence 20 AA;

Query Match 22.1%; Score 32.5; DB 21; Length 20;

Best Local Similarity 52.6%; Pred. No. 2e+02; Mismatches 1; Gaps 1;

Matches 10; Conservative 1; Indels 7; Indels 1; Gaps 1;

QY 1 LCKAFNSTLP-TMAQMEKA 18

I: |||: | | | |

Db 2 lakvfsgsdipdmfcmka 20

RESULT 11

AAG83226

ID AAG83226 standard; Peptide; 20 AA.

XX AC AAG83226;

XX 05-SEP-2001 (first entry)

XX Chlamydia Swib peptide 61-80.

XX Chlamydia; vaccine; infection; fusion protein; antigen;

KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;

KW acute respiratory tract infection; Cap1; CT529; OMCB;

KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.

XX Chlamydia sp.

OS WO200140474-A2.

PN 07-JUN-2001.

PD 04-DEC-2000; 2000WO-US32919.

XX 03-DEC-1999; 99US-0454684.

PR 19-APR-2000; 2000US-0556877.

PR 20-JUN-2000; 2000US-0598419.

XX (CORI-) CORIXA CORP.

PA Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;

PI WPI; 2001-374831/39.

XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic

PT inflammatory disease, trachoma, acute respiratory tract infections,

PT atherosclerosis and heart disease -

XX Claim 66; Page 235; 295pp; English.

PS The present peptide is provided in a specification relating to

XX compounds and methods for the treatment and diagnosis of chlamydial

CC infection. The compounds provided include polypeptides and fusion

CC proteins comprising immunogenic portions of Chlamydia antigens

CC and DNA sequences encoding such polypeptides. They are useful for

CC vaccinating against chlamydial infection, which causes pelvic

CC inflammatory disease, trachoma, acute respiratory tract infections,

CC atherosclerosis and heart disease.

XX Sequence 20 AA;

XX

Query Match 22.1%; Score 32.5; DB 22; Length 20;

Best Local Similarity 52.8%; Pred. No. 2e-02;

Matches 10; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 1 LCRAFNSTLP-TMAQMEKA 18

Db 2 lakvfssdpdmfqtka 20

RESULT 12

AAE03101

ID AAE03101 standard; Protein; 20 AA.

XX AAE03101;

XX 10-AUG-2001 (first entry)

XX Human gene 13 encoded secreted protein HOEDE28, SEQ ID NO:135.

XX Human; secreted protein; proliferative disorder; cancer; tumour;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW inflammation; allergy; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

KW cardiovascular disorder; angiotensin-related disorder; kidney disorder;

KW gastrointestinal disorder; pregnancy-related disorder; gene therapy;

KW endocrine disorder; infection; wound healing; vulnery;

KW cell culture; chemotaxis; food additive; chromosome 15;

KW binding partner identification.

XX Homo sapiens.

XX Location/Qualifiers

FT Key 1..16

FT Peptide /label= signal\_peptide

FT Protein 17..20

FT /note= "Mature secreted protein"

XX WO200132676-A1.

PN 10-MAY-2001.

XX 25-OCT-2000; 2000WO-US29365.

XX 29-OCT-1999; 99US-0162237.

PR 21-JUL-2000; 2000US-0219666.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Shi Y, Olsen HS, Soppet DR;

PI WPI; 2001-328773/34.

XX N-PSDB; AAD07620.

XX Nucleic acids encoding 25 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. Gaucher's disease,

PT Alzheimer's disease, Scimitar syndrome, Creutzfeldt-Jacob disease,

PT diabetes mellitus and multiple sclerosis -

XX Claim 11; Page 482; 546pp; English.

XX AAD07571-AAD07645 represent cDNAs corresponding to 25 human secreted

CC protein genes, and AAE03052-AAE03126 represent the proteins they encode.

CC AAD03127-AAE03150 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing

CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the

CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the

CC 25 genes, based on the tissues in which they are most highly expressed,

CC and include developing products for the diagnosis or treatment of

CC proliferative disorders, cancer, tumours, foetal and developmental

CC abnormalities, haematopoietic disorders, diseases of the immune system,

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

CC allergies, neurological disorders (e.g., Alzheimer's disease,

CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,

CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

CC cardiovascular disorders, angiogenic disorders, kidney disorders,

CC gastrointestinal disorders, pregnancy-related disorders, endocrine

CC disorders, and infections. The proteins can also be used to aid wound

CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell

CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties.

CC Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked

CC immunosorbent assay (ELISA). The present sequence represents a human

CC secreted protein of the invention.

XX Sequence 20 AA;

XX

Query Match 21.8%; Score 32; DB 22; Length 20;

Best Local Similarity 66.7%; Pred. No. 2.4e-02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 ALSIGFETC 26

Db 3 alsisfctc 11



## RESULT 13

AA13659  
ID AAB13659 standard; Peptide; 22 AA.

XX  
AC  
AA13659;

02-FEB-2001 (first entry)

DE C. pneumoniae Swib peptide 66-87.

XX Chlamydia pneumoniae.  
KW Chlamydial infection; sexually transmitted disease;  
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
KW trachoma; blindness; acute respiratory tract infection;  
KW atherosclerosis; coronary heart disease; antibacterial.

XX Chlamydia pneumoniae.

PN WO200034483-A2.

PD 15-JUN-2000.

PF 08-DEC-1999; 99WO-US29012.

PR 08-DEC-1998; 98US-0208277.

PR 08-APR-1999; 99US-0288594.

PR 01-OCT-1999; 99US-0410568.

PR 22-OCT-1999; 99US-0426571.

XX (CORI-) CORIXA CORP.

PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;

XX WPI; 2000-431303/37.

XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises  
PT amino acid sequence encoded by polynucleotide sequence -

PS Disclosure; Page 229; 256pp; English.

XX The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and can  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a protein isolated in the  
CC present invention.

XX Sequence 22 AA;

Query Match 21.4%; Score 31.5; DB 21; Length 22;

Best Local Similarity 58.8%; Pred. No. 3.2e+02;

Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 5 FNSTLP-TMAQMEKALS 20

Db 1 fgssdpdmfqtkaals 17

||:| | |||||

## RESULT 14

AA13659  
ID AAG83227 standard; Peptide; 22 AA.

XX

AC AAG83227;

XX 05-SEP-2001 (first entry)

XX Chlamydia Swib peptide 66-87.

XX Chlamydia; vaccine; infection; fusion protein; antigen;

KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;

KW acute respiratory tract infection; Cap1; C529; OMCB;

KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.

XX Chlamydia sp.

PN WO200140474-A2.

XX 07-JUN-2001.

XX 04-DEC-2000; 2000WO-US32919.

XX 03-DEC-1999; 99US-0454684.

PR 19-APR-2000; 2000US-0556877.

PR 20-JUN-2000; 2000US-0598419.

XX (CORI-) CORIXA CORP.

PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;

XX WPI; 2001-374831/39.

XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic  
PT inflammatory disease, trachoma, acute respiratory tract infections,  
PT atherosclerosis and heart disease -

XX Claim 66; Page 235; 295pp; English.

XX The present peptide is provided in a specification relating to  
CC compounds and methods for the treatment and diagnosis of chlamydial  
CC infection. The compounds provided include polypeptides and fusion  
CC proteins comprising immunogenic portions of Chlamydia antigens  
CC and DNA sequences encoding such polypeptides. They are useful for  
CC vaccinating against chlamydial infection, which causes pelvic  
CC inflammatory disease, trachoma, acute respiratory tract infections,  
CC atherosclerosis and heart disease.

XX Sequence 22 AA;

Query Match 21.4%; Score 31.5; DB 22; Length 22;

Best Local Similarity 58.8%; Pred. No. 3.2e+02;

Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 5 FNSTLP-TMAQMEKALS 20

Db 1 fgssdpdmfqtkaals 17

||:| | |||||

## RESULT 15

AA13659

ID AAY19410 standard; Peptide; 18 AA.

XX

AC AAY19410;

XX 14-JUL-1999 (first entry)

XX Lecithin:cholesterol acyltransferase activation exhibiting peptide #223.

XX Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;

KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;

KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;

KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;

XX septic shock.

XX Synthetic.

```

OS Homo sapiens.
XX
PN WO9916459-A1.
XX
PD 08-APR-1999.
XX
PF 28-SEP-1998; 98WO-US20327.
XX
PR 29-SEP-1997; 97US-0940095.
XX
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (DASS/) DASSEUX J.
PA (DUFO/) DUFOURCQ J.
PA (METZ/) METZ G.
PA (SEKU/) SEKUL R.
XX
PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;
PI Sekul R;
XX
DR WPI; 1999-277035/23.
XX
PT Peptide agonists of apolipoprotein A-I
XX
PS Example; Page 128; 280pp; English.
XX
CC The present invention describes an agonist (A) of apolipoprotein A-I
CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC complexes, are used to treat or prevent diseases associated with
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular
CC disease, atherosclerosis, restenosis, HDL (high density lipoprotein) or
CC apoA-I deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC treating septic shock. When labeled, (A) can also be used diagnostically
CC to measure serum levels of HDL, in particular the HDL subpopulation that
CC is involved in retrograde cholesterol transport, also to image HDL at
CC e.g. atherosclerotic streaks, and to raise antibodies. AAY19188 to
CC AAY19441 represent lecithin:cholesterol acyltransferase (LCAT) activity
CC exhibiting core peptides, which are apoA-I agonists.
XX
SQ Sequence 18 AA;

```

Query Match 21.1%; Score 31; DB 20; Length 18;  
 Best Local Similarity 42.9%; Pred. No. 3e+02;  
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

QY 4 AFNSTLPTMAQMEK 17
   ||| :||:|
Db 5 afrellealqlkk 18

```

Search completed: February 21, 2002, 17:12:17  
 Job time: 193 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:10:44 ; Search time 12.47 Seconds  
(without alignments)  
50.529 Million cell updates/sec

Title: US-08-753-851-13  
Perfect score: 147  
Sequence: 1 LCKAFNSTLPTMAQMEKALSIGFETCRY 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 129629

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	28	2 US-08-143-311B-11	Sequence 11, Appl
2	34	23.1	26	2 US-08-751-767A-67	Sequence 67, Appl
3	33	22.4	28	1 US-07-977-630-45	Sequence 45, Appl
4	31	21.1	18	3 US-08-940-095-223	Sequence 223, App
5	31	21.1	18	3 US-08-940-093-223	Sequence 223, App
6	31	21.1	18	3 US-08-940-096-223	Sequence 223, App
7	31	21.1	18	4 US-09-465-719-223	Sequence 223, App
8	30.5	20.7	21	6 5164481-9	Patent No. 5164481
9	30	20.4	15	6 5169933-30	Patent No. 5169933
10	30	20.4	16	2 US-08-817-933A-7	Sequence 7, Appl
11	30	20.4	16	4 US-08-602-999A-226	Sequence 226, App
12	30	20.4	17	5 PCT-US95-02121-96	Sequence 96, Appl
13	30	20.4	18	2 US-09-024-198-15	Sequence 15, Appl
14	30	20.4	18	2 US-09-186-409-15	Sequence 15, Appl
15	30	20.4	19	2 US-09-024-198-1	Sequence 1, Appl
16	30	20.4	19	2 US-09-186-409-1	Sequence 1, Appl
17	30	20.4	21	1 US-07-908-455A-6	Sequence 6, Appl
18	30	20.4	21	1 US-08-434-120-32	Sequence 32, Appl
19	30	20.4	21	1 US-08-186-266-6	Sequence 6, Appl
20	30	20.4	21	1 US-08-465-325-32	Sequence 32, Appl
21	30	20.4	21	1 US-08-446-692-48	Sequence 48, Appl
22	30	20.4	21	2 US-08-488-351A-48	Sequence 48, Appl
23	30	20.4	21	3 US-09-100-409A-54	Sequence 54, Appl
24	30	20.4	21	5 PCT-US95-02121-97	Sequence 27, Appl
25	30	20.4	21	5 PCT-US95-13841-20	Sequence 20, Appl
26	30	20.4	22	2 US-08-143-311B-1	Sequence 1, Appl
27	30	20.4	24	1 US-08-385-375-14	Sequence 14, Appl

Sequence 15, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 17, Appl  
Sequence 8, Appl  
Sequence 7, Appl  
Sequence 33, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 9, Appl  
Sequence 9, Appl  
Sequence 7, Appl  
Sequence 478, App  
Sequence 355, App  
Sequence 26, Appl

28 30 20.4 26 1 US-08-385-375-15  
29 30 20.4 28 1 US-08-214-770-7  
30 30 20.4 28 5 PCT-US95-02885-7  
31 30 20.4 28 5 PCT-US95-07542-17  
32 29 19.7 12 2 US-09-057-762-8  
33 29 19.7 12 3 US-08-326-119A-8  
34 29 19.7 21 1 US-07-908-455A-7  
35 29 19.7 21 1 US-08-434-120-33  
36 29 19.7 21 1 US-08-465-325-33  
37 29 19.7 27 3 US-08-809-397-23  
38 29 19.7 27 5 PCT-US95-12502-23  
39 28.5 19.4 17 2 US-08-509-187D-9  
40 28.5 19.4 17 5 PCT-US93-09704A-9  
41 28.5 19.4 26 2 US-08-637-759B-478  
42 28.5 19.4 26 3 US-08-871-355A-478  
43 28 19.0 9 4 US-09-258-754-355  
44 28 19.0 9 4 US-09-042-107-355  
45 28 19.0 15 1 US-08-268-251-26

## ALIGNMENTS

RESULT 1  
US-08-143-311B-11  
; Sequence 11, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143.311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-11

; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,

ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 009196-0004-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 223:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6004925e  
US-08-940-093-223

Query Match 21.1%; Score 31; DB 3; Length 18;  
Best Local Similarity 42.9%; Pred. No. 77;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 AFNSTLPTMAQMEK 17  
|| | :||:|  
Db 5 AFRELLEAQLKK 18

RESULT 5  
US-08-940-093-223  
; Sequence 223, Application US/08940093  
; Patent No. 6037323  
; GENERAL INFORMATION:  
; APPLICANT: Dasseux, Jean-Louis  
; APPLICANT: Sekul, Renate  
; APPLICANT: Buttner, Klaus  
; APPLICANT: Cornut, Isabelle  
; APPLICANT: Metz, Gunther  
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS  
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
; NUMBER OF SEQUENCES: 258  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/940,093  
; FILING DATE: 29-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 009196-0006-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEFAX: 650-493-5556  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 223:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: No. 6037323e  
US-08-940-093-223

Query Match 21.1%; Score 31; DB 3; Length 18;  
Best Local Similarity 42.9%; Pred. No. 77;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 AFNSTLPTMAQMEK 17  
|| | :||:|  
Db 5 AFRELLEAQLKK 18

RESULT 6  
US-08-940-096-223  
; Sequence 223, Application US/08940096  
; Patent No. 6046166  
; GENERAL INFORMATION:  
; APPLICANT: Dasseux, Jean-Louis  
; APPLICANT: Sekul, Renate  
; APPLICANT: Buttner, Klaus  
; APPLICANT: Cornut, Isabelle  
; APPLICANT: Metz, Gunther  
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS  
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
; NUMBER OF SEQUENCES: 258  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/940,096  
; FILING DATE: 29-SEP-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 009196-0005-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEFAX: 650-493-5556  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 223:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6046166e  
US-08-940-096-223

Query Match 21.1%; Score 31; DB 3; Length 18;  
Best Local Similarity 42.9%; Pred. No. 77;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 AFNSTLPTMAQMEK 17  
|| | :||:|  
Db 5 AFRELLEAQLKK 18

RESULT 7

US-09-465-719-223  
 ; Sequence 223, Application US/09465719  
 ; Patent No. 6265377  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dasseux, Jean-Louis  
 ; APPLICANT: Sekul, Renate  
 ; APPLICANT: Buttner, Klaus  
 ; APPLICANT: Cornut, Isabelle  
 ; APPLICANT: Metz, Gunther  
 ; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS  
 ; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS  
 ; NUMBER OF SEQUENCES: 258  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10036-2811  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/465,719  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/940,093  
 ; FILING DATE: 29-SEP-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 009196-0006-999  
 ; TELEPHONE: 650-493-4935  
 ; TELEFAX: 650-493-5556  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 223:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 18 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: No. 6265377e  
 ; US-09-465-719-223

Query Match 21.1%; Score 31; DB 4; Length 18;  
 Best Local Similarity 42.9%; Pred. No. 77;  
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 AFNSTLPTMAQMEK 17  
 II I :||:|  
 Db 5 AFRELLEAQLKK 18

RESULT 8  
 5164481-9  
 ; Patent No. 5164481  
 ; APPLICANT: LACROIX, MARTIAL; REIN, MAAN; DIONNE, GERVAIS  
 ; TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES  
 ; FOR DETECTING AND ELICITING ANTIBODIES TO RUBELLA VIRUS  
 ; NUMBER OF SEQUENCES: 11  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/397,767  
 ; FILING DATE: 23-AUG-1989  
 ; SEQ ID NO: 9:  
 ; LENGTH: 21  
 5164481-9

Query Match 20.7%; Score 30.5; DB 6; Length 21;  
 Best Local Similarity 57.1%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 7 STLP-TMAQMEKAL 19  
 ||| ||| :|||  
 Db 3 STTPTMEDLQAL 16

RESULT 9  
 5169933-30  
 ; Patent No. 5169933  
 ; APPLICANT: ANDERSON, DAVID C.; MORGAN, CHARLES JR.; FRITZBERG,  
 ; ALAN R.; NICHOLS, EVERETT J.  
 ; TITLE OF INVENTION: CAVALENTLY-LINKED COMPLEXES AND METHODS  
 ; FOR ENHANCED CYTOTOXICITY AND IMAGING  
 ; NUMBER OF SEQUENCES: 45  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/390,241  
 ; FILING DATE: 07-AUG-1989  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 232,337  
 ; FILING DATE: 15-AUG-1988  
 ; SEQ ID NO: 30:  
 ; LENGTH: 15  
 5169933-30

Query Match 20.4%; Score 30; DB 6; Length 15;  
 Best Local Similarity 60.0%; Pred. No. 90;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 12 MAQMEKALSI 21  
 :||:|||||:  
 Db 3 IAKMEKASSV 12

RESULT 10  
 US-08-817-933A-7  
 ; Sequence 7, Application US/08817933A  
 ; Patent No. 5945104  
 ; GENERAL INFORMATION:  
 ; APPLICANT: STANWORTH, DENIS R.  
 ; APPLICANT: LEWIN, IAN V.  
 ; TITLE OF INVENTION: PEPTIDES FOR ANTI-ALLERGY TREATMENT  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Nixon & Vanderhye  
 ; STREET: 1100 No. 5945104th Glebe Road, 8th Floor  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: US  
 ; ZIP: 22201-4714  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/817,933A  
 ; FILING DATE: 27-JUN-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9422294.0  
 ; FILING DATE: 04-NOV-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mitchard, Leonard C.  
 ; REGISTRATION NUMBER: 29,009  
 ; REFERENCE/DOCKET NUMBER: 179-23  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 816-4005  
 ; TELEFAX: (703) 816-4100  
 ; TELEX: N/A

INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-817-933A-7

Query Match 20.4%; Score 30; DB 2; Length 16;  
Best Local Similarity 60.0%; Pred. No. 98;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 12 MAOMEKALSI 21  
Db 4 IAKMEKASSV 13

RESULT 11  
US-08-602-999A-226  
Sequence 226, Application US/08602999A  
Patent No. 618-205  
GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILLIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999A  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 226:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-602-999A-226

Query Match 20.4%; Score 30; DB 4; Length 16;  
Best Local Similarity 42.9%; Pred. No. 98;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 FNSTLPTMAQMEKA 18  
I: ||| :||

Db 2 FDFPLPTHPNLQKA 15

RESULT 12  
PCT-US95-02121-96  
Sequence 96, Application PC/TUS9502121  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
TITLE OF INVENTION: CTL IMMUNITY  
NUMBER OF SEQUENCES: 153  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02121  
FILING DATE: 16-FEB-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/197,484  
FILING DATE: 16-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1-17  
OTHER INFORMATION: /note="Malaria circumsporozoite  
OTHER INFORMATION: 382-398"  
PCT-US95-02121-96

Query Match 20.4%; Score 30; DB 5; Length 17;  
Best Local Similarity 60.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 12 MAOMEKALSI 21  
Db 2 IAKMEKASSV 11

RESULT 13  
US-09-024-198-15  
Sequence 15, Application US/09024198  
Patent No. 5912323  
GENERAL INFORMATION:  
APPLICANT: FASANO, Alessio  
TITLE OF INVENTION: ZONULA OCCLUDENS TOXIN RECEPTOR

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 19 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

RESULT 14  
US-09-186-409-15  
; Sequence 15, Application US/09186409  
; Patent No. 5948629  
; GENERAL INFORMATION:  
; APPLICANT: FASANO, Alessio  
; TITLE OF INVENTION: ZONULA OCCLUDENS TOXIN RECEPTOR  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
; CITY: Washington, D.C.  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/186,409  
; FILING DATE: 5 NOV 1998  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 09/024,198



US-09-024-198-1  
HYPOTHETICAL: NO

Query Match 20.4%; Score 30; DB 2; Length 19;  
Best Local Similarity 45.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 MAQMEKALSIG 22  
: : : : :  
Db 2 LTELEKALNXG 12

Search completed: February 21, 2002, 17:12:36  
Job time: 112 sec



Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 9 HVVIPRIHP 17  
| : | : |  
Db 6 HNVLPNIEP 14

## RESULT 3

B49254  
Tcr C gamma 1 chain V-J region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Jul-1999  
C:Accession: B49254  
R:Esquerre, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E.M.  
Eur. J. Immunol. 22, 491-498, 1992  
A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of the T  
A:Reference number: A49037; MUID:92164730  
A:Accession: B49254  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-18 <EQ>  
A:Cross-references: GB:S90688; NID:g246308; PIDN:RAB21557.1; PID:g246309  
A:Experimental source: Y93A cells  
A:Note: sequence extracted from NCBI backbone (NCBIN:90688, NCBIIP:90690)  
C:Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 23.0%; Score 28; DB 2; Length 18;  
Best Local Similarity 50.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CRYGFIEG 8  
| : | : |  
Db 5 CSYGSSG 12

## RESULT 4

C56049  
superoxide dismutase (EC 1.15.1.1) (Fe/Mn) [similarity] - unidentified organism (fragment)  
C:Species: unidentified organism  
C:Date: 12-Apr-1995 #sequence\_revision 12-Apr-1995 #text\_change 20-Apr-2000  
C:Accession: C56049  
R:Binette, J.P.; Binette, M.B.  
Scanning Microsc. 8, 233-239, 1994  
A:Title: Sequencing of proteins extracted from stones.  
A:Reference number: A56049; MUID:95215817  
A:Accession: C56049  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-19 <BIN>  
A:Experimental source: urate-calcium oxalate kidney stones  
A:Note: the source is designated as Homo sapiens, however the true source probably origi  
C:Superfamily: superoxide dismutase (Mn)  
C:Keywords: metalloprotein; oxidoreductase

Query Match 23.0%; Score 28; DB 2; Length 19;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YGFIEGHV 10  
| : | : |  
Db 12 YGALEXHI 19

## RESULT 5

B41868  
hypothetical protein (traE1 3' region) - Enterococcus faecalis plasmid pAD1  
C:Species: Enterococcus faecalis  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999  
C:Accession: B41868; B37391  
R:Pontius, L.T.; Clewell, D.B.  
J. Bacteriol. 174, 3152-3160, 1992

A:Title: Conjugative transfer of Enterococcus faecalis plasmid pAD1: nucleotide sequ  
A:Reference number: A41868; MUID:92250408  
A:Contents: plasmid pAD1  
A:Accession: B41868  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-15 <PON>  
A:Note: sequence extracted from NCBI backbone (NCBIN:99901, NCBIIP:99906)  
R:Clewell, D.B.; Pontius, L.T.; An, F.Y.; Ike, Y.; Suzuki, A.; Nakayama, J.  
Plasmid 24, 156-161, 1990  
A:Title: Nucleotide sequence of the sex pheromone inhibitor (iAD1) determinant of Ent  
A:Reference number: A37391; MUID:91261999  
A:Accession: B37391  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-15 <CLE>  
A:Cross-references: GB:M62888; NID:g141853; PIDN:AAA98040.1; PID:g141855  
C:Genetics:  
A:Genome: plasmid

Query Match 22.1%; Score 27; DB 2; Length 15;  
Best Local Similarity 83.3%; Pred. No. 4.3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 HVVIPR 14  
| : | : | : |  
Db 5 HVVIPR 10

## RESULT 6

F57789  
gallbladder stone matrix protein 2, 41K - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 03-May-1996  
C:Accession: F57789  
R:Binette, J.P.; Binette, M.B.  
Submitted to the Protein Sequence Database, February 1996  
A:Description: The proteins of gallbladder stones.  
A:Reference number: A57789  
A:Accession: F57789  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <BIN>  
A:Note: 9-Phe was also found

Query Match 22.1%; Score 27; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GFIEG 8  
| : | : | : |  
Db 7 GFIEG 11

## RESULT 7

H35141  
T-cell receptor delta chain V region (105.23) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Aug-1990 #sequence\_revision 31-Aug-1990 #text\_change 30-May-1997  
C:Accession: H35141  
R:Sim, G.K.; Augustin, A.  
Cell 61, 397-405, 1990  
A:Title: Dominantly inherited expression of BID, an invariant undiversified T cell re  
A:Reference number: A35141; MUID:90242386  
A:Accession: H35141  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-16 <SIM>  
C:Keywords: T-cell receptor

Query Match 21.3%; Score 26; DB 2; Length 16;  
Best Local Similarity 38.5%; Pred. No. 6.7e+02;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CRYGFIEGHVVP 13  
I I I I I  
Db 1 CASGYILGGIRAP 13

RESULT 8  
S12904  
protein kinase (EC 2.7.1.37) - starfish (Pisaster ochraceus)  
C:Species: Pisaster ochraceus  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Aug-1997  
C:Accession: S12904  
R:Sanghera, J.S.; Abersold, R.; Morrison, H.D.; Bures, E.J.; Pelech, S.L.  
FEBS Lett. 273, 223-226, 1990  
A:Title: Identification of the sites in myelin basic protein that are phosphorylated by  
A:Reference number: S12904; MUID:91032186  
A:Accession: S12904  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <SAS>  
C:Keywords: phosphotransferase

Query Match 20.5%; Score 25; DB 2; Length 14;  
Best Local Similarity 45.5%; Pred. No. 8.5e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 9 HVVIPRIHPNS 19  
: I I I I I  
Db 1 NIVTPRTPPPS 11

RESULT 9  
A36133  
hypothetical protein (bkd 5' region) - Pseudomonas putida (fragment)  
C:Species: Pseudomonas putida  
C:Date: 30-Nov-1990 #sequence\_revision 30-Nov-1990 #text\_change 30-Sep-1993  
C:Accession: A36133  
R:Madhusudan, K.T.; Huang, G.; Burns, G.; Sokatch, J.R.  
J. Bacteriol. 172, 5655-5663, 1990  
A:Title: Transcriptional analysis of the promoter region of the Pseudomonas putida brand  
A:Reference number: A36133; MUID:91008935  
A:Accession: A36133  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-18 <MAD>  
A:Cross-references: GB:M33715

Query Match 20.5%; Score 25; DB 2; Length 18;  
Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 GFIEG 8  
I I I I I  
Db 5 GFLEG 9

RESULT 10  
S02643  
RNA-directed RNA polymerase (EC 2.7.7.48) - tobacco mosaic virus (fragment)  
N:Alternate names: RNA replicase  
C:Species: tobacco mosaic virus, TMV  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 30-Sep-1993  
C:Accession: S02643  
R:Weshi, T.; Motoyoshi, F.; Adachi, A.; Watanabe, Y.; Takamatsu, N.; Okada, Y.  
EMBO J. 7, 1575-1581, 1988  
A:Title: Two concomitant base substitutions in the putative replicase genes of tobacco m  
A:Reference number: S02643  
A:Accession: S02643

Query Match 20.1%; Score 24.5; DB 2; Length 21;  
Best Local Similarity 60.0%; Pred. No. 1.6e+03;  
Matches 6; Conservative 1; Mismatches 0; Indels 3; Gaps 1;  
QY 2 RYGFIEGHV 11  
I I I I I  
Db 7 RY---EGHVM 13

RESULT 11  
A26999  
carboxylesterase (EC 3.1.1.1), intestinal - Caenorhabditis elegans (fragment)  
C:Species: Caenorhabditis elegans  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 07-Feb-1997  
C:Accession: A26999  
R:McGhee, J.D.  
Biochemistry 26, 4101-4107, 1987  
A:Title: Purification and characterization of a carboxylesterase from the intestine o  
A:Reference number: A26999; MUID:88000636  
A:Accession: A26999  
A:Molecule type: protein  
A:Residues: 1-13 <MCG>  
C:Keywords: carboxylic ester hydrolase; intestine

Query Match 19.7%; Score 24; DB 2; Length 13;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 YGFIEG 8  
I I I I I  
Db 8 YGKVEG 13

RESULT 12  
A47393  
neuropeptide callatostatin 1 - bluebottle fly (Calliphora vomitoria)  
N:Contains: neuropeptide callatostatin 2  
C:Species: Calliphora vomitoria  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47393; B47393; C47393  
R:Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993  
A:Title: Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with seq  
A:Reference number: A47393; MUID:93211980  
A:Accession: A47393  
A:Molecule type: protein  
A:Residues: 1-16 <DUV>  
A:Experimental source: thoracic ganglia, brains, heads  
A:Note: sequence extracted from NCBI backbone (NCBIP:128476)  
A:Accession: B47393  
A:Molecule type: protein  
A:Residues: 3-16 <DU2>  
A:Experimental source: head  
A:Note: sequence extracted from NCBI backbone (NCBIP:128477)  
A:Accession: C47393  
A:Molecule type: protein  
A:Residues: 1-8 <DU3>  
A:Experimental source: thoracic ganglia  
A:Note: sequence extracted from NCBI backbone (NCBIP:128478)  
C:Function:  
A:Description: inhibits juvenile hormone production  
C:Keywords: amidated carboxyl end; neuropeptide  
F:1-16/Product: callatostatin 1 #status experimental <MA1>

F:3-16/Product: callatostatin 2 #status experimental <NA2>  
F:9-16/Product: callatostatin 3 #status experimental <NA3>  
F:16/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 19.7%; Score 24; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYGF 5  
|||  
Db 11 RYGF 14

## RESULT 13

B86323  
protein F14D16.5 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: B86323  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;  
anssen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, X.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, Y.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: B86323  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-17 <STO>  
A:Cross-references: GB:AF005172; NID:g8778270; PIDN:AAF79279.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F14D16.5  
A:Map position: 1

Query Match 19.7%; Score 24; DB 2; Length 17;  
Best Local Similarity 25.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 IECHVVIPIHPNSIC 21  
::: |||  
Db 1 MDAHKAHLSPPSRC 16

## RESULT 14

DRDPPP  
distal-retinal-pigment hormone - northern shrimp  
N:Alternate names: light-adapting hormone  
C:Species: Pandalus borealis (northern shrimp)  
C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 23-Aug-1996  
C:Accession: A01473  
R:fernund, P.  
Biochim. Biophys. Acta 439, 17-25, 1976  
A:Title: Structure of a light-adapting hormone from the shrimp, Pandalus borealis.  
A:Reference number: A01473; MUID:76253762  
A:Accession: A01473  
A:Molecule type: protein  
A:Residues: 1-18 <FER>  
C:Comment: This peptide causes migration of the distal retinal pigment into the proximal  
the amount of light entering the retinulas.  
C:Superfamily: pigment-dispersing hormone  
C:Keywords: amidated carboxyl end; hormone  
F:18/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 19.7%; Score 24; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 1.6e+03;

Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 4 GFIEGHVIVPRI 15  
|||  
Db 3 GMINSILGIPRV 14

## RESULT 15

I39461  
anti-angiotensin, hypothetical - human (fragment)  
N:Alternate names: hypothetical angiotensin receptor antagonist, angiotensin mRNA com  
C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000  
C:Accession: I39461  
R:Moore, G.J.; Ganter, R.C.; Franklin, K.J.  
Biochem. Biophys. Res. Commun. 160, 1387-1391, 1989  
A:Title: Angiotensin 'antipeptides': (-)-messenger RNA complementary to human angioten  
A:Reference number: I39461; MUID:89273605  
A:Accession: I39461  
A:Molecule type: mRNA  
A:Residues: 1-18 <MOO>  
A:Cross-references: GB:M26228; NID:gl78641; PIDN:AAA35530.1; PID:gl78642  
A:Note: this sequence is the conceptual translation of an nucleotide sequence complem

Query Match 19.7%; Score 24; DB 4; Length 18;  
Best Local Similarity 42.9%; Pred. No. 1.6e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 IHPNSIC 21  
|||  
Db 12 VHPVTCC 18

Search completed: February 21, 2002, 16:44:17  
Job time: 191 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:52:21 ; Search time 20.15 Seconds  
(without alignments)  
38.211 Million cell updates/sec

Title: US-08-753-851-14  
Perfect score: 122  
Sequence: 1 CRYGFEIGHVVPRIHPNSIC 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 1160

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	19.7	16	1 ALL1_CALVO	P41839 Calliphora
2	24	19.7	18	1 DRPH_PANBO	P01209 pandalus bo
3	24	19.7	21	1 DCMS_PSECA	P19921 pseudomonas
4	23	18.9	19	1 UKAI_HUMAN	P31940 homo sapien
5	22	18.0	12	1 PVK2_PERAM	P81555 periplaneta
6	22	18.0	16	1 ARCD_PSEPU	P41147 pseudomonas
7	22	18.0	18	1 AROF_STRRM	P80575 streptomyces
8	21.5	17.6	15	1 RM12_YEAST	P36522 saccharomyc
9	21	17.2	10	1 ANGT_CHICK	P01018 gallus gall
10	21	17.2	10	1 BPPB_BOTIN	P30426 bothrops in
11	21	17.2	10	1 GON3_ONCKE	P20367 oncorhynch
12	21	17.2	11	1 TKN3_PSEGU	P42988 pseudophryn
13	21	17.2	13	1 LMT4_LOCMI	P41490 locusta mig
14	21	17.2	19	1 DUCR_STRGP	P36503 streptomyces
15	21	17.2	20	1 FRHA_METBA	P80489 methanosarc
16	20	16.4	9	1 NEUX_HUMAN	P04277 homo sapien
17	20	16.4	13	1 TAL3_TREME	P01370 tremella me
18	20	16.4	15	1 UC08_MAIZE	P80614 zea mays (m
19	20	16.4	19	1 FIBB_PIG	P14477 sus scrofa
20	19	15.6	8	1 ANG2_BOTJA	Q10582 bothrops ja
21	19	15.6	9	1 NEUX_RAT	P11382 rattus norv
22	19	15.6	10	1 GLEM_HUMAN	P02728 homo sapien
23	19	15.6	14	1 ANGT_HORSE	P01016 equus cabal
24	19	15.6	19	1 CAT3_FASHE	P80532 fasciola he
25	19	15.6	19	1 CXA2_CONST	P28879 conus stria
26	19	15.6	19	1 LPRM_STAOU	P03063 staphylococ
27	19	15.6	20	1 KORA_METTM	P80904 salmonobact
28	19	15.6	20	1 OSMB_SALTY	P37723 salmonella
29	19	15.6	21	1 PSAL_SYNVU	P25937 synecococc
30	19	15.6	21	1 CSPS_STRTR	P81622 streptococc
31	18	14.8	9	1 PPK1_PERAM	P82691 periplaneta
32	18	14.8	9	1 TAL3_PICJA	P17441 pichia jadi
33	18	14.8	10	1 AH3_PRUSE	P29261 prunus sero

34 18 14.8 10 1 ANG1\_BOTJA Q10581 bothrops ja  
35 18 14.8 10 1 ANGT\_BOVIN P01017 bos taurus  
36 18 14.8 11 1 ANGT\_CRIGE P09037 crinia geor  
37 18 14.8 11 1 BPPB\_AKHA P01021 agkistrodon  
38 18 14.8 11 1 MLG\_THETS P41989 theromyzon  
39 18 14.8 11 1 TIN1\_HOPTI P82651 hoplobatrach  
40 18 14.8 13 1 AH4\_PRUSE P29262 prunus sero  
41 18 14.8 13 1 TPI3\_PHYRO P04096 phyllomedus  
42 18 14.8 14 1 COCO\_LIMPO P35866 limulus pol  
43 18 14.8 14 1 DCMW\_PSECF P19914 pseudomonas  
44 18 14.8 14 1 HY14\_PIG P01155 sus scrofa  
45 18 14.8 15 1 DCMW\_PSECH P19917 pseudomonas

## ALIGNMENTS

RESULT 1  
ALL1\_CALVO  
ID ALL1\_CALVO STANDARD; PRT; 16 AA.  
AC P41839;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CALLATOSTATIN 1 (LEU-CALLATOSTATIN 1) [CONTAINS: CALLATOSTATIN 2 (LEU-CALLATOSTATIN 2); CALLATOSTATIN 3 (LEU-CALLATOSTATIN 3)].  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_taxonomy:27454;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Thoracic ganglion, Brain, and Head;  
RX MEDLINE=93211980; PubMed=8460157;  
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S., Thorpe A.;  
RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with sequence homology to cockroach allatostatins";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=94291167; PubMed=8020069;  
RA Duve H., Thorpe A.;  
RT "Distribution and functional significance of Leu-callatostatins in the blowfly Calliphora vomitoria";  
RL Cell tissue Res. 276:367-379(1994).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.  
CC -!- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEBRAL COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY SYSTEM AND INTESTINE.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
DR PIR: A47393; A47393.  
DR PIR: B47393; B47393.  
DR PIR: C47393; C47393.  
KW Neuropeptide; Amidation.  
FT PEPTIDE 1 16 CALLATOSTATIN 1.  
FT PEPTIDE 3 16 CALLATOSTATIN 2.  
FT PEPTIDE 9 16 CALLATOSTATIN 3.  
FT MOD\_RES 16 16 AMIDATION.  
SQ SEQUENCE 16 AA; 1907 MW; A435B68C26EC3D09 CRC64;

Query Match 19.7%; Score 24; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYGF 5  
||||

Db 11 RYGF 14

```

RESULT 2
DRPH_PANBO STANDARD; PRT; 18 AA.
AC P01209;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PIGMENT-DISPERSING HORMONE (PDH) (LIGHT ADAPTING DISTAL RETINAL
DE PIGMENT HORMONE) (DRPH).
OS Pandalus borealis (Northern red shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidae;
OC Pandalidae; Pandalus.
OX NCBI_TaxID=6703;
RN [1]
RP SEQUENCE.
RX MEDLINE=76253762; PubMed=952951;
RA Fernlund P.;
RT "Structure of a light-adapting hormone from the shrimp, Pandalus
RT borealis."
RL Biochim. Biophys. Acta 439:17-25(1976).
CC -1- FUNCTION: CAUSES THE MIGRATION OF THE DISTAL RETINAL PIGMENT
CC INTO THE PROXIMAL END OF THE PIGMENT CHROMATOPHORE CELLS AND
CC THUS DECREASES THE AMOUNT OF LIGHT ENTERING THE RETINULAS.
CC -1- SIMILARITY: TO THE PDH OF OTHER ARTHROPODS.
DR PIR; A01473; DRDPPP.
KW Hormone; Amidation.
FT MOD_RES 18 18 AMIDATION.
SQ SEQUENCE 18 AA; 1903 MW; B2346B6D0178650E CRC64;

Query Match 19.7%; Score 24; DB 1; Length 18;
Best Local Similarity 41.7%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 GFTEGHVVIPRI 15
| | : | | :
DB 3 GMINSILGIPRV 14

RESULT 3
DCMS_PSECA STANDARD; PRT; 21 AA.
AC P19921;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] SMALL CHAIN (EC 1.2.2.4)
DE {FRAGMENT}.
OS Pseudomonas carboxydovorans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Oligotropha.
OX NCBI_TaxID=40137;
RN [1]
RP SEQUENCE.
RX STRAIN=OM5;
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxyotrophic bacteria."
RL Arch. Microbiol. 152:335-341(1989).
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 = CO(2) +
CC 2 H(+) + FERRICYTOCHROME B-561.
CC -1- COFACTOR: MOLYBDENUM.
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P01144;
KW Oxidoreductase; Molybdenum.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2270 MW; 68D4380629401B9C CRC64;

```

```

Query Match 19.7%; Score 24; DB 1; Length 21;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 6 IEGHVVPRIHP 17
| | | : |
DB 10 INGHPEALVEP 21

RESULT 4
UKAJ_HUMAN STANDARD; PRT; 19 AA.
ID UKAJ_HUMAN
AC P31940;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF EPIDERMAL KERATINOCYTES (SPOT 1118)
DE (FRAGMENTS).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX TISSUE=Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes."
RL Electrophoresis 13:960-969(1992).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 7.24, ITS MW IS: 23.5 KDA.
DR Aarhus/Ghent-2DPAGE; 1118; IEF.
FT NON_TER 1 1
FT UNSURE 6 6
FT NON_CONS 6 7
FT NON_CONS 12 13
FT UNSURE 17 17
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2087 MW; EF7515F79D50DE12 CRC64;

Query Match 18.9%; Score 23; DB 1; Length 19;
Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 HWVVPRIHPNSI 20
| : | : | :
DB 1 HGLVRLTPTEV 12

RESULT 5
PVK2_PERAM STANDARD; PRT; 12 AA.
ID PVK2_PERAM
AC P81555;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PERIVISCEROKININ-2 (PEA-PVK-2).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattodea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RX TISSUE=Abdominal perisymphathetic organs;
RX MEDLINE=98326577; PubMed=9663444;
RA Predel R., Rapus J., Eckert M., Holman G.M., Nachman R.J., Wang Y.,
RA Penzlin H.;

```



RT "Isolation of periviscerokinin-2 from the abdominal perisymphathetic  
 RL organs of the American cockroach, *Periplaneta americana*."  
 CC Peptides 19:801-809(1998).  
 CC -|- FUNCTION: MYOACTIVE PEPTIDE; HAS EXCITATORY ACTIONS ON THE  
 CC HYPERNEURAL MUSCLES.  
 CC -|- MASS SPECTROMETRY: MW=1189.3; METHOD=MALDI.  
 KW Neuropeptide; Amidation.  
 FT MOD.RES 12 12 AMIDATION  
 SQ SEQUENCE 12 AA; 1190 MW; 2F4D8EE1EB05728 CRC64;

Query Match 18.0%; Score 22; DB 1; Length 12;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 GFIEGHVPIRI 15  
 | | | : ||:  
 Db 1 GSSSGLISMPRV 12

RESULT 6  
 ARCD\_PSEPU STANDARD; PRT; 16 AA.  
 AC P41147;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ARGININE/ORNITHINE ANTIporter (FRAGMENT).  
 GN ARCD.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 4359;  
 RA Wilson S.D., Wang M., Filpula D.;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 CC -|- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGinine  
 CC AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE  
 CC ARGinine DEIMINASE PATHWAY.  
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
 CC (POTENTIAL).  
 CC -|- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF  
 CC PERMEASES.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; 007185; AAA16963.1;  
 KW Transport; Amino-acid transport; Transmembrane; Inner membrane.  
 FT NON\_TER 1  
 SQ SEQUENCE 16 AA; 1644 MW; 90B48A7C8FAA9705 CRC64;

Query Match 18.0%; Score 22; DB 1; Length 16;  
 Best Local Similarity 30.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGFIEGHVVI 12  
 | | | : :  
 Db 7 YGLYDGLTL 16

RESULT 7  
 AROF\_STRRM STANDARD; PRT; 18 AA.  
 AC P80575;

DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PHOSPHO-2-DEHYDRO-3-DEOXYHEPTONATE ALDOLASE (EC 4.1.2.15) (PHOSPHO-2-  
 DE KETO-3-DEOXYHEPTONATE ALDOLASE) (DAHP SYNTHETASE) (3-DEOXY-D-ARABINO-  
 DE HEPTULOSONATE 7-PHOSPHATE SYNTHASE) (FRAGMENT).  
 GN AROH.  
 OS Streptomyces rimosus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1927;  
 RN [1]  
 RP SEQUENCE AND CHARACTERIZATION.  
 RC STRAIN-ISOLATE 4018;  
 RX MEDLINE=96349103; PubMed=8760910;  
 RA Walker G.E., Dunbar B., Hunter I.S., Nimmo H.G., Coggins J.R.;  
 RT "Evidence for a novel class of microbial  
 RT 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase in Streptomyces  
 RT coelicolor A3(2). Streptomyces rimosus and Neurospora crassa.";  
 RL Microbiology 142:1973-1982(1996).  
 CC -|- CATALYTIC ACTIVITY: 7-PHOSPHO-2-DEHYDRO-3-DEOXY-D-ARABINO-  
 CC HEPTONATE + ORTHOPHOSPHATE = PHOSPHOENOLPYRUVATE + D-ERYTHROSE  
 CC 4-PHOSPHATE + H(2)O.  
 CC -|- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN  
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
 CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -|- PTM: THE N-TERMINUS IS BLOCKED.  
 CC -|- SIMILARITY: BELONGS TO CLASS-II DAHP SYNTHETASE FAMILY.  
 DR InterPro: IPR002480; DAHP\_synth\_2.  
 DR Pfam: PF01474; DAHP\_synth\_2; 1.  
 DR Aromatic amino acid biosynthesis; Lyase.  
 KW NON\_TER 1  
 FT NON\_TER 18  
 SQ SEQUENCE 18 AA; 2209 MW; 7FE627E713C871BB CRC64;

Query Match 18.0%; Score 22; DB 1; Length 18;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+03;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GFIEGH 9  
 | | | |  
 Db 13 GFEVH 18

RESULT 8  
 RM12\_YEAST STANDARD; PRT; 15 AA.  
 ID RM12\_YEAST  
 AC P36522;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L12 (YML12) (FRAGMENT).  
 GN MRPL12.

OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=91285106; PubMed=2060626;  
 RA Grohmann L., Graack H.-R., Kruft V., Choli T., Goldschmidt-Reisin S.,  
 RA Kitakawa M.;  
 RT "Extended N-terminal sequencing of proteins of the large ribosomal  
 RT subunit from yeast mitochondria.";  
 RL FEBS Lett. 284:51-56(1991).  
 DR PIR; S17261; S17261.  
 DR SGD; L0002687; MRPL12.

KW Ribosomal protein; Mitochondrion.  
 FT NON\_TER 15  
 SQ SEQUENCE 15 AA; 1851 MW; 74BCD9FEDDB3900 CRC64;

Query Match 17.6%; Score 21.5; DB 1; Length 15;  
 Best Local Similarity 46.2%; Pred. No. 2.1e+03;  
 Matches 6; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 5 FIEGHVVIPIHP 17  
 Db 6 FVE---VIVRFNP 15  
 1:1 11111

RESULT 9  
 ANGT\_CHICK STANDARD; PRT; 10 AA.

AC FO1018;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ANGIOTENSINOGEN [CONTAINS: ANGIOTENSIN I; ANGIOTENSIN II] (FRAGMENT).  
 GN SERPINAB OR AGI.  
 OS Gallus gallus (Chicken), and  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031, 93934;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=Chicken; PubMed=4361802;  
 RX MEDLINE=74127845; PubMed=4361802;  
 RA Nakayama T., Nakajima T., Sokabe H.;  
 RT "Comparative studies on angiotensins. 3. Structure of fowl  
 angiotensin and its identification by DNS-method.";  
 RL Chem. Pharm. Bull. 21:2085-2087(1973).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=C.c.japonica;  
 RX MEDLINE=90284684; PubMed=2191893;  
 RA Takei Y., Hasegawa Y.;  
 RT "Vasopressor and depressor effects of native angiotensins and  
 inhibition of these effects in the Japanese quail.";  
 RL Gen. Comp. Endocrinol. 79:12-22(1990).  
 CC -1- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN  
 CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN  
 CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE  
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT  
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL  
 CC BALANCE OF BODY FLUIDS.  
 CC -1- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR PIR; A01250; A01250.  
 DR PIR; A90917; A90917.  
 DR PIR; A60624; A60624.  
 DR InterPro: IPR000215; Serpin.  
 DR PROSITE: PS00284; SERPIN; PARTIAL.  
 KW Vasoconstrictor; Plasma; Serpin.  
 FT PEPTIDE 1 10 ANGIOTENSIN I.  
 FT PEPTIDE 1 8 ANGIOTENSIN II.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1232 MW; CEPBDD761F2DB42 CRC64;

Query Match 17.2%; Score 21; DB 1; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 15 IHPNSI 20  
 Db 5 VHPFSL 10  
 1:1 11

RESULT 10  
 BPP8\_BOTIN STANDARD; PRT; 10 AA.  
 ID BPP8\_BOTIN  
 AC P30426;

DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE BRADYKININ-POTENTIATING PEPTIDE S5,1 (ANGIOTENSIN-CONVERTING  
 ENZYME INHIBITOR).  
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 OC Viperidae; Crotalinae; Bothrops.  
 OX NCBI\_TaxID=8723;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Venom;  
 RX MEDLINE=90351557; PubMed=2386615;  
 RA Cintra A.C.O., Vieira C.A., Giglio J.R.;  
 RT "Primary structure and biological activity of bradykinin potentiating  
 peptides from Bothrops insularis snake venom.";  
 RL J. Protein Chem. 9:221-227(1990).  
 CC -1- FUNCTION: THIS PEPTIDE BOTH ENHANCES THE ACTIVITY OF THE  
 ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF  
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.  
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.  
 DR PIR; H37196; H37196.  
 KW Hypotensive agent; Venom.  
 FT MOD\_RES 1 1  
 SQ SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 17.2%; Score 21; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 HPN 18  
 Db 5 HPN 7  
 111

RESULT 11  
 GON3\_ONCKE STANDARD; PRT; 10 AA.

ID GON3\_ONCKE  
 AC P20367; P81751;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE GONADOLIBERIN III (GONADOTROPIN-RELEASING HORMONE III) (GNRH-III) (LH-  
 RH III) (LULIBERIN III).  
 GN GNRH3.  
 OS Oncorhynchus keta (Chum salmon), and  
 OS Clupea pallasii (Pacific herring).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8018, 30724;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=O.keta;  
 RX MEDLINE=83195140; PubMed=6341999;  
 RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;  
 RT "Characterization of a teleost gonadotropin-releasing hormone.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).  
 RN [2]  
 RP SEQUENCE, AND FUNCTION.  
 RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;  
 RX MEDLINE=20114351; PubMed=10650929;  
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,  
 RA Chang J.P., Rivier J.E., Sherwood N.M.;  
 RT "Primary structure and function of three gonadotropin-releasing  
 hormones, including a novel form, from an ancient teleost, herring.";  
 RL Endocrinology 141:505-512(2000).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 DR PIR; A21114; A21114.

DR InterPro: IPR002012; GnRH.  
 DR Pfam: PF00446; GnRH; 1.  
 DR PROSITE: PS00473; GnRH; 1.  
 KW Hormone; Amidation; Hypothalamus.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 17.2%; Score 21; DB 1; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 YGFIEG 8  
 ||::|  
 DB 5 YGWLPG 10

RESULT 12  
 TKN3\_PSEGU STANDARD; PRT; 11 AA.  
 AC P42988;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PSEUDOPHYRNE GUENTHERI K-III (PG-KIII).  
 OS Pseudophryne guentheri (Frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
 OC Pseudophryne.  
 OX NCBI\_TaxID=30349;  
 RN [1]  
 RP SEQUENCE.

RC TISSUE-Skin;  
 RX MEDLINE=90287814; PubMed=2356157;  
 RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,  
 RA Roberts J.D., Melchiorri P., Espamer V.;  
 RT "Six novel tachykinin- and bombesin-related peptides from the skin of  
 RT the Australian frog Pseudophryne guntheri.";  
 RL Peptides 11:299-304(1990).  
 CC -|- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.

-|- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.

DR PIR: D60409; D60409.  
 DR InterPro: IPR003580; Protachykinin.  
 DR Pfam: PF02202; Tachykinin.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1268 MW; 3DBA7C37C9CB1457 CRC64;

Query Match 17.2%; Score 21; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 HPN 18  
 ||||  
 DB 3 HPN 5

RESULT 13  
 LMT4\_LOCFI STANDARD; PRT; 13 AA.  
 ID LMT4\_LOCFI  
 AC P41490;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE LOCUSTAMYOTROPIN 4 (LOM-MT-4).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE-Brain;  
 RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,  
 RA de Loof A.;  
 RT "Isolation, identification and synthesis of locustamyotropin III and  
 RT IV, two additional neuropeptides of Locusta migratoria: members of the  
 RT locustamyotropin peptide family.";  
 RL Insect Biochem. Mol. Biol. 22:447-452(1992).  
 CC -|- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTILE ACTIVITY  
 CC (MYOTROPIC ACTIVITY). LOM-MT IV SEEMS TO BE A MORE POTENT MUSCLE  
 CC STIMULATOR THAN LOM-MT I, II AND III.  
 CC -|- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 DR InterPro: IPR001484; Pyrokinin.  
 DR PROSITE: PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin.  
 FT MOD\_RES 13 13 AMIDATION.  
 SQ SEQUENCE 13 AA; 1553 MW; 20861943824D6698 CRC64;

Query Match 17.2%; Score 21; DB 1; Length 13;  
 Best Local Similarity 42.9%; Pred. No. 2.2e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 14 RIHPNSI 20  
 ||::|  
 DB 1 RLHONGM 7

RESULT 14  
 DURC\_STRGP STANDARD; PRT; 19 AA.  
 AC P36503;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE LANTIBIOTIC DURAMYCIN C.  
 OS Streptomyces griseolutes.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=29306;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=R2107;  
 RX MEDLINE=91107436; PubMed=2125590;  
 RA Fredenhagen A., Fendrich G., Marki F., Marki W., Gruner J.,  
 RA Raschdorf F., Peter H.H.;  
 RA "Duramycins B and C, two new lanthionine containing antibiotics as  
 RT inhibitors of phospholipase A2. Structural revision of duramycin and  
 RT cinnamycin";  
 RL J. Antibiot. 43:1403-1412(1990).  
 RN [2]  
 RP STRUCTURE BY NMR.  
 RA Zimmermann N., Freund S., Fredenhagen A., Jung G.;  
 RT "Solution structure of the lantibiotics duramycin B and C.";  
 RL (In) Schneider C.H., Eberles A.N. (eds.);  
 RL Peptides 1992, pp.519-520, Escom Science Publishers, Leiden (1993).  
 RN [3]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=93387292; PubMed=8375380;  
 RA Zimmermann N., Freund S., Fredenhagen A., Jung G.;  
 RT "Solution structures of the lantibiotics duramycin B and C.";  
 RL Eur. J. Biochem. 216:419-428(1993).  
 CC -|- FUNCTION: ACTS AS INHIBITOR OF PHOSPHOLIPASE A2.  
 CC -|- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF  
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE  
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND

Search completed: February 21, 2002, 16:52:21  
Job time: 520 sec

CC CLEAVAGE OF THE MODIFIED PRECURSOR.  
KW -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE B LANTIBIOTICS.  
KW Antibiotic; Bacteriocin; Lantibiotic.  
FT MOD\_RES 4 4 D-ALANINE.  
FT MOD\_RES 6 6 D-ALANINE (BONDED TO 19).  
FT MOD\_RES 11 11 D-ABU (AMINOBTYRIC ACID).  
FT MOD\_RES 18 18 D-ABU (AMINOBTYRIC ACID).  
FT MOD\_RES 19 19 LYSINOALANINE (BONDED TO 6).  
FT THIOETH 1 18 ABU-S-CYS (METHYLLANTHIONINE).  
FT THIOETH 4 14 ALA-S-CYS (LANTHIONINE).  
FT THIOETH 5 11 ABU-S-CYS (METHYLLANTHIONINE).  
SQ SEQUENCE 19 AA; 2007 MW; E2404ECE3F95286A CRC64;

Query Match 17.2%; Score 21; DB 1; Length 19;

Best Local Similarity 75.0%; Pred. No. 3.2e+03;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRYG 4

Db 5 CSYG 8

RESULT 15  
FRHA\_METBA STANDARD; PRT; 20 AA.  
AC P80489;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE COENZYME F420 HYDROGENASE ALPHA SUBUNIT (EC 1.12.99.1) (8-HYDROXY-5-DEAZAFLAVIN-REDUCING HYDROGENASE ALPHA SUBUNIT) (FRH) (FRAGMENT).  
GN FRHA.  
OS Methanosarcina barkeri.  
OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;  
OC Methanosarcina.  
OX NCBI\_TaxID=2208;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=FUSARO / DSM 804;  
RX MEDLINE=96085134; PubMed=8521835;  
RA Michel R., Massanz C., Kostka S., Richter M., Fiebig K.  
RT "Biochemical characterization of the 8-hydroxy-5-deazaflavin-reactive hydrogenase from Methanosarcina barkeri Fusaro."  
RL Eur. J. Biochem. 233:727-735(1995).  
CC -1- FUNCTION: REDUCES THE PHYSIOLOGICAL LOW-POTENTIAL TWO-ELECTRON ACCEPTOR COENZYME F420, AND THE ARTIFICIAL ONE-ELECTRON ACCEPTOR METHYLVIIOLOGEN.  
CC -1- CATALYTIC ACTIVITY: H(2) + COENZYME F420 = REDUCED COENZYME F420.  
CC -1- COFACTOR: FRH CONTAINS NICKEL, IRON-SULFUR, AND FAD COFACTORS.  
CC -1- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS AND A GAMMA CHAIN.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.  
CC -1- SIMILARITY: BELONGS TO THE [NIFE]/[NIFESE] HYDROGENASE LARGE SUBUNIT FAMILY.  
DR InterPro: IPR001501; Nifese\_Hases.  
DR PROSITE: PS00507; NI\_HGENASE\_L\_1; PARTIAL.  
DR PROSITE: PS00508; NI\_HGENASE\_L\_2; PARTIAL.  
KW Oxidoreductase; Nickel; FAD.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2233 MW; 0D659741D776D03 CRC64;

Query Match 17.2%; Score 21; DB 1; Length 20;

Best Local Similarity 75.0%; Pred. No. 3.3e+03;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 IEHG 9

Db 12 LEHG 15



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:51:55 ; Search time 60.03 Seconds  
(without alignments)  
51.170 Million cell updates/sec

Title: US-08-753-851-14

Perfect score: 122

Sequence: 1 CRVGFIEGHVVIPRIHNSIC 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 5819

Minimum DB seq length: 0

Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- SPTREMBL\_17.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	63	51.6	12	4 Q9UC29	Q9uc29 homo sapien
2	38	31.1	17	4 Q9UC28	Q9uc28 homo sapien
3	31	25.4	11	5 Q9TWM2	Q9twm2 alypsia cal
4	30	24.6	16	2 Q44610	Q44610 buchnera ap
5	29	23.8	20	10 Q9X1W8	Q9x1w8 oryza sativ
6	29	23.8	21	2 Q44611	Q44611 buchnera ap
7	28.5	23.4	19	12 Q83273	Q83273 cucumber mo
8	28	23.0	18	4 Q9UCF4	Q9ucf4 homo sapien
9	28	23.0	19	4 Q9UC82	Q9uc82 homo sapien
10	27	22.1	15	2 Q52135	Q52135 enterococcu
11	27	22.1	17	7 P79480	P79480 cervus elap
12	27	22.1	20	5 Q9TWM9	Q9twm9 octopus vul
13	27	22.1	20	11 Q9QVB2	Q9qvb2 mus sp. ser
14	27	22.1	21	12 Q9PXB3	Q9pxb3 duck hepati
15	26	21.3	9	8 P92072	P92072 euhadra her
16	26	21.3	19	5 P82600	P82600 aedes aegypt
17	26	21.3	20	2 Q9RSV7	Q9rsv7 pseudomonas
18	25	20.5	17	4 Q9UCN0	Q9ucn0 homo sapien
19	25	20.5	20	6 Q9TRM7	Q9trm7 bos taurus

20	24	19.7	10	13	Q9PRY8	Q9pry8 triakis scy
21	24	19.7	17	2	Q9R4C4	Q9r4c4 agrobacteri
22	24	19.7	17	4	Q95795	Q95795 homo sapien
23	24	19.7	17	10	Q9LMD7	Q9lmd7 arabidopsis
24	24	19.7	17	13	P82396	P82396 littoriana
25	24	19.7	18	4	Q13767	Q13767 homo sapien
26	24	19.7	20	12	Q41541	Q41541 human immun
27	23	18.9	14	6	Q09061	Q09061 bos taurus
28	23	18.9	15	4	Q9UCB9	Q9ucb9 homo sapien
29	23	18.9	18	4	Q9UCB8	Q9ucb8 homo sapien
30	22.5	18.4	14	13	Q91777	Q91777 xenopus lae
31	22	18.0	10	4	Q9UCQ4	Q9ucq4 homo sapien
32	22	18.0	10	12	Q75595	Q75595 human immun
33	22	18.0	11	6	Q9XSP8	Q9xsp8 presbytis j
34	22	18.0	11	6	Q9XSP7	Q9xsp7 pygathrix n
35	22	18.0	11	6	Q9XSP6	Q9xsp6 pongo pygma
36	22	18.0	11	6	Q9XSP2	Q9xsp2 hylobates s
37	22	18.0	11	11	Q61797	Q61797 mus musculus
38	22	18.0	12	12	Q88575	Q88575 theiler's e
39	22	18.0	12	12	Q88576	Q88576 theiler's e
40	22	18.0	12	12	Q88577	Q88577 theiler's e
41	22	18.0	12	12	Q88578	Q88578 theiler's e
42	22	18.0	12	12	Q88579	Q88579 theiler's e
43	22	18.0	12	12	Q88580	Q88580 theiler's e
44	22	18.0	12	12	Q88581	Q88581 theiler's e
45	22	18.0	12	12	Q88582	Q88582 theiler's e

ALIGNMENTS

RESULT 1

Q9UC29  
ID Q9UC29 PRELIMINARY; PRT; 12 AA.  
AC Q9UC29;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE 85 KDA GLIOMA MEMBRANE PROTEIN/CD44 HOMOLOG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95079448; PubMed=7527301;  
RA Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;  
RT "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes  
RT overexpressed CD44 in glioma cells."  
RL Cancer Immunol. Immunother. 39:313-317(1994).  
SQ SEQUENCE 12 AA; 1337 MW; 2E0F6CE9D9D2C1E8 CRC64;

Query Match 51.6%; Score 63; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GFIEGHVVIPRI 15

Db 1 GFIEGHVVIPRI 12

RESULT 2

Q9UC28  
ID Q9UC28 PRELIMINARY; PRT; 17 AA.  
AC Q9UC28;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE 85 KDA GLIOMA MEMBRANE PROTEIN/CD44 HOMOLOG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]  
RP SEQUENCE  
RX MEDLINE=95079448; PubMed=7527301;  
RA Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;  
RT "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes  
overexpressed CD44 in glioma cells."  
RL Cancer Immunol. Immunother. 39:313-317(1994).  
SQ SEQUENCE 17 AA; 1788 MW; 7806F51C97A3E217 CRC64;

Query Match 31.1%; Score 38; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 HPNSIC 21

Db 1 HPNSIC 6

RESULT 3

Q9TWM2  
ID Q9TWM2 PRELIMINARY; PRT; 11 AA.  
AC Q9TWM2  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE BUCCALIN B, BUCH.  
OS Aplysia californica (California sea hare).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anasipidea;  
OC Aplysiidae; Aplysia.  
OX NCBI\_TaxID=6500;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95083478; PubMed=7991459;  
RA Villm F.S., Cropper E.C., Rosen S.C., Tenenbaum R., Kupfermann I.,  
RA Weiss K.R.;  
RT "Structure, localization, and action of buccalin B: a bioactive  
peptide from Aplysia."  
RL Peptides 15:959-969(1994).  
SQ SEQUENCE 11 AA; 1153 MW; 692253F9C9C86B44 CRC64;

Query Match 25.4%; Score 31; DB 5; Length 11;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 RYGFIEG 8

Db 4 RYGFVGG 10

RESULT 4

Q44610  
ID Q44610 PRELIMINARY; PRT; 16 AA.  
AC Q44610;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE SHIKIMATE DEHYDROGENASE (FRAGMENT).  
GN AROE.  
OS Buchnera aphidicola.  
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
OX NCBI\_TaxID=9;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95212914; PubMed=7535281;  
RA Rouhakhsh D., Baumann P.;  
RT "Characterization of a putative 23S-5S rRNA operon of Buchnera  
aphidicola (endosymbiont of aphids) unlinked to the 16S rRNA-encoding  
gene."  
RL Gene 155:107-112(1995).  
DR EMBL; U10499; AAA79128.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 21 AA; 2614 MW; 1A1B28C9E9D8007A CRC64;

Query Match 23.8%; Score 29; DB 2; Length 21;  
Best Local Similarity 55.6%; Pred. No. 8.6e+02;

FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1891 MW; 72A8175598D30DF1 CRC64;

Query Match 24.6%; Score 30; DB 2; Length 16;  
Best Local Similarity 38.3%; Pred. No. 4.6e+02;  
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 FIEGHVVIPRIHP 17

Db 2 FLEFHGIFPKIEP 14

RESULT 5

Q9XIW8  
ID Q9XIW8 PRELIMINARY; PRT; 20 AA.  
AC Q9XIW8;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE HYPOTHETICAL PROTEIN.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartioideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Nagamura Y., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC  
clone:P0681F10."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB026295; BAA81863.1; -.  
SQ SEQUENCE 20 AA; 2312 MW; A5EE9C9093D159F9 CRC64;

Query Match 23.8%; Score 29; DB 10; Length 20;  
Best Local Similarity 71.4%; Pred. No. 8.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 14 RIHPNSI 20

Db 6 RIHDNSV 12

RESULT 6

Q44611  
ID Q44611 PRELIMINARY; PRT; 21 AA.  
AC Q44611;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE SHIKIMATE DEHYDROGENASE (FRAGMENT).  
GN AROE.  
OS Buchnera aphidicola.  
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
OX NCBI\_TaxID=9;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95212914; PubMed=7535281;  
RA Rouhakhsh D., Baumann P.;  
RT "Characterization of a putative 23S-5S rRNA operon of Buchnera  
aphidicola (endosymbiont of aphids) unlinked to the 16S rRNA-encoding  
gene."  
RL Gene 155:107-112(1995).  
DR EMBL; U10500; AAA79129.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 21 AA; 2614 MW; 1A1B28C9E9D8007A CRC64;

Query Match 23.8%; Score 29; DB 2; Length 21;  
Best Local Similarity 55.6%; Pred. No. 8.6e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 HVVIPRIHP 17  
 Db 6 HNVLPNIEP 14

RESULT 7  
 Q83273 PRELIMINARY; PRT; 19 AA.  
 AC Q83273;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE ORF II.  
 OS cucumber mosaic virus (cucumber mosaic cucumovirus).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;  
 OC Cucumovirus.  
 OX NCBI\_TaxID=12305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T43;  
 RA Masuta C., Hayashi Y., Wang W.Q., Takanami Y.;  
 RT "Comparison of four satellite RNA isolates of cucumber mosaic virus."  
 RL Ann. Phytopathol. Soc. Jpn. 56:207-212(1990).  
 DR EMBL: D10039; BAA00929.1; -.  
 SQ SEQUENCE 19 AA; 2467 MW; 51A5A51EDDD2F38B CRC64;

Query Match 23.4%; Score 28.5; DB 12; Length 19;  
 Best Local Similarity 58.3%; Pred. No. 9.3e+02;  
 Matches 7; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 CRY-GFIEGHVV 11  
 Db 3 CRYHGFKKHSV 14

RESULT 8  
 Q9UCF4 PRELIMINARY; PRT; 18 AA.  
 AC Q9UCF4;  
 DT 01-WAY-2000 (TREMBLrel. 13, Created)  
 DT 01-WAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE CHYMOTRYPSINOGEN HOMOLOG (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93285747; PubMed=8509158;  
 RA Kim Y.J., Zhou Z., Hurtado J., Wood D.L., Choi A.S., Pescovitz M.D.,  
 RA Warfel K.A., Vandagriff J., Davis J.K., Kwon B.S.;  
 RT "IDM patients' sera recognize a novel 30-kD pancreatic autoantigen  
 RT related to chymotrypsinogen."  
 RL Immunol. Invest. 22:219-227(1993).  
 SQ SEQUENCE 18 AA; 1826 MW; 38105D1037CA6837 CRC64;

Query Match 23.0%; Score 28; DB 4; Length 18;  
 Best Local Similarity 66.7%; Pred. No. 1e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 IPRIPHP 17  
 Db 3 VPAIRP 8

RESULT 9  
 Q9UC82 PRELIMINARY; PRT; 19 AA.  
 ID Q9UC82

Q9UC82;  
 AC 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE 22 KDA STONE MATRIX PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95215817; PubMed=7701298;  
 RT Binette J.P., Binette M.B.;  
 RT "Sequencing of proteins extracted from stones."  
 RL Scanning Microsc. 8:233-239(1994).  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR HSSP; P17670; LIDS.  
 DR InterPro; IPR001189; SOD\_M1.  
 DR Pfam; PF00081; sodfe; 1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 19 AA; 2162 MW; AD703A074F3C9655 CRC64;

Query Match 23.0%; Score 28; DB 4; Length 19;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 YGFIEGHV 10  
 Db 12 YGALEXHI 19

RESULT 10  
 Q52135 PRELIMINARY; PRT; 15 AA.  
 ID Q52135;  
 AC Q52135;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE PLASMID PAD1 SEX PHEROMONE INHIBITOR (IAD1) DETERMINANT.  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Plasmid PAD1.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
 OC Enterococcus.  
 OX NCBI\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91261999; PubMed=2128961;  
 RA Clewell D.B., Pontius L.T., An F.Y., Ike Y., Suzuki A., Nakayama J.;  
 RT "Nucleotide sequence of the sex pheromone inhibitor (IAD1) determinant  
 RT of Enterococcus faecalis conjugative plasmid PAD1."  
 RL Plasmid 24:156-161(1990).  
 DR EMBL: M62888; AAA98040.1; -.  
 KW Plasmid.  
 SQ SEQUENCE 15 AA; 1874 MW; 0D9D07E3079E3559 CRC64;

Query Match 22.1%; Score 27; DB 2; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 HVVIPR 14  
 Db 5 HVVIPR 10

RESULT 11  
 P79480 PRELIMINARY; PRT; 17 AA.  
 ID P79480



```

AC P79480;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DE MHC CLASS II DRB (FRAGMENT).
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860;
RN [1]
RP SEQUENCE FROM N.A.
RA Swabrick P.A., Crawford A.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U63083; AAB37783.1; -.
KW MHC.
FT NON_TER 1 1
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 2022 MW; 2B7EEBCF17F004FA CRC64;

Query Match 22.1%; Score 27; DB 7; Length 17;
Best Local Similarity 53.8%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 1 CR--YGFIEGHV 11
   || || ||
DB 1 CRHNYGVIESFTV 13

RESULT 12
Q9TWM9
ID Q9TWM9 PRELIMINARY; PRT; 20 AA.
AC Q9TWM9;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE ALPHAM, ALPHA-MACROGLOBULIN PROTEINASE INHIBITOR (FRAGMENT).
OS Octopus vulgaris (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=6645;
RN [1]
RP SEQUENCE.
RA Thøgersen I.B., Salvesen G., Brucato F.H., Pizzo S.V., Enghild J.J.;
RL "Purification and characterization of an alpha-macroglobulin
proteinase inhibitor from the mollusc Octopus vulgaris.";
RL Biochem. J. 285:521-527(1992).
SQ SEQUENCE 20 AA; 2370 MW; 4A903B6BBF226BD9 CRC64;

Query Match 22.1%; Score 27; DB 5; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 2; Gaps 1;

QY 7 EGH--VVIPR 14
   ||| :|||
DB 5 EGHYMLVPR 14

RESULT 13
Q9QVB2
ID Q9QVB2 PRELIMINARY; PRT; 20 AA.
AC Q9QVB2;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TremBLrel. 13, Last annotation update)
DE SERINE PROTEASE (FRAGMENT).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;

```

```

RN [1]
RP SEQUENCE.
RX MEDLINE=93020730; PubMed=1404084;
RA Damjanov A., Damjanov I.;
RT "Isolation of serine protease from granulated metrial gland cells of
mice and rats with lectin from Dolichos biflorus.";
RL J. Reprod. Fertil. 95:679-684(1992).
SQ SEQUENCE 20 AA; 2290 MW; 0A304F61A22C500D CRC64;

Query Match 22.1%; Score 27; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 8 GHVVIPIHP 17
   ||| |||
DB 4 GHEVPHSRP 13

RESULT 14
Q9PXB3
ID Q9PXB3 PRELIMINARY; PRT; 21 AA.
AC Q9PXB3;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE 170 KDA DHEV PRE-S REGION BINDING PROTEIN (FRAGMENT).
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE.
RX MEDLINE=96013813; PubMed=7474130;
RA Tong S., Li J., Wands J.R.;
RT "Interaction between duck hepatitis B virus and a 170-kilodalton
cellular protein is mediated through a neutralizing epitope of the
pre-S region and occurs during viral infection.";
RL J. Virol. 69:7106-7112(1995).
SQ SEQUENCE 21 AA; 2329 MW; A464BC8B8F8F230E CRC64;

Query Match 22.1%; Score 27; DB 12; Length 21;
Best Local Similarity 33.3%; Pred. No. 1.8e+03;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 10 VVIPRIHPN 18
   |::| :|:
DB 8 VIVPSLNPD 16

RESULT 15
P92072
ID P92072 PRELIMINARY; PRT; 9 AA.
AC P92072;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DT 01-AUG-1998 (TremBLrel. 07, Last annotation update)
DE ATPASE SUBUNIT 8 (FRAGMENT).
OS Euhadra herklotsi.
OC Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Helicacea; Bradybaenidae; Euhadra.
OX NCBI_TaxID=58912;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEPATOPANCREAS;
RA Yamazaki N., Ueshima R., Terrett J.A., Yokobori S., Kaifu M.,
RA Segawa R., Kobayashi T., Numachi K., Ueda T., Nishikawa K.,
RA Watanabe K., Thomas R.H.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z71697; CAA96373.1; -.
KW Mitochondrion.
ET NON_TER 9 9

```

SQ SEQUENCE 9 AA; 977 MW; 25BBB1F775B736C7 CRC64;

Query Match 21.3%; Score 26; DB 8; Length 9;  
Best Local Similarity 33.3%; Pred. No. 4.7e+05;  
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 12 IPRHPNSI 20  
: : : : :  
Db 1 VPQLSPHSL 9

Search completed: February 21, 2002, 16:51:55  
Job time: 534 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:42:58 ; Search time 64.88 Seconds  
(without alignments)  
23.976 Million cell updates/sec

Title: US-08-753-851-14  
Perfect score: 122  
Sequence: 1 CRYGFIEGHVVPRIHPNSIC 21

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 206698

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	21	15 AAR53484	CD44 peptide CD44-
2	104	85.2	19	15 AAR53485	CD44 peptide CD44-
3	32.5	26.6	16	19 AAW81195	Synthetic human nH
4	32	26.2	9	22 AAU03735	Cyclic peptide inh
5	32	26.2	9	22 AAU03751	Cyclic peptide inh
6	32	26.2	12	18 AAW36883	HTLV-I protease cl
7	32	26.2	18	18 AAW26612	Chick muscle-speci
8	31	25.4	9	20 AAW83548	Tumour necrosis fa
9	31	25.4	10	20 AAW83549	Tumour necrosis fa
10	31	25.4	11	20 AAW83550	Tumour necrosis fa
11	31	25.4	12	20 AAW83551	Tumour necrosis fa

12	31	25.4	13	20 AAW83552	Tumour necrosis fa
13	31	25.4	14	20 AAW83553	Tumour necrosis fa
14	31	25.4	15	20 AAW83554	Tumour necrosis fa
15	31	25.4	16	18 AAW35207	Diastereomer pep
16	31	25.4	16	19 AAW82923	Antipathogenic pep
17	31	25.4	16	20 AAW83555	Tumour necrosis fa
18	31	25.4	16	21 AAB17458	Antipathogenic pep
19	31	25.4	17	20 AAW83556	Tumour necrosis fa
20	31	25.4	18	20 AAW83557	Tumour necrosis fa
21	30	24.6	10	15 AAY38164	Hepatitis B virus-
22	30	24.6	10	20 AAY45732	Immunogenic peptid
23	30	24.6	10	21 AAB23536	Angiogenic vascula
24	30	24.6	10	22 AAG97961	Human complementar
25	30	24.6	10	22 AAG97962	Human complementar
26	30	24.6	13	20 AAY48698	Membrane dipeptida
27	30	24.6	14	15 AAR53696	Japanese cedar pol
28	30	24.6	15	17 AAR97942	Japan cedar pollen
29	30	24.6	15	17 AAR97943	Japan cedar pollen
30	30	24.6	15	21 AAY57842	Human leptin recep
31	30	24.6	18	22 AAM19282	Peptide #5716 enco
32	30	24.6	18	22 AAM32006	Peptide #6043 enco
33	30	24.6	20	14 AAR54374	RAE 27.1. Ambrosi
34	30	24.6	20	17 AAW02458	RAE 27.1 comprisin
35	30	24.6	21	20 AAY15089	Human protease-act
36	29	23.8	12	21 AAY93878	Reactive peptide w
37	29	23.8	15	10 AAP90252	Antigenic peptide
38	29	23.8	15	20 AAY25576	Human MHC Class II
39	29	23.8	15	20 AAY25577	Human MHC Class II
40	29	23.8	15	20 AAY25575	Human MHC Class II
41	29	23.8	17	20 AAY08385	Cysteine noose lib
42	29	23.8	19	22 AAB76504	Peptide which bind
43	29	23.8	20	18 AAW22108	Leptin receptor an
44	29	23.8	21	20 AAY25560	Human MHC Class II
45	29	23.8	21	22 AAM17964	Peptide #4398 enco

ALIGNMENTS

RESULT 1  
AAR53484  
ID AAR53484 standard; peptide; 21 AA.  
XX AC AAR53484;  
XX AC AAR53484;  
DT 01-DEC-1994 (first entry)  
XX CD44 peptide CD44-11.  
XX DE CD44 peptide CD44-11.  
XX KW Cell adhesion molecule; CD44; antinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autoimmune disease;  
KW immunosuppressive.  
XX OS Homo sapiens.  
XX PN WO9409811-A.  
XX PD 11-MAY-1994.  
XX PF 29-OCT-1993; 93WO-US10412.  
XX PR 30-OCT-1992; 92US-0973339.  
XX (UYDU-) UNIV DUKE.  
XX PI Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
XX DR WPI; 1994-167121/20.  
XX Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis

CONFIDENTIAL

DE Cyclic peptide inhibitor of IFA-1/TCAM-1 interaction #12. XX

XX Cyclic; lymphocyte function associated antigen-1; LFA-1; asthma;  
 KW intracellular adhesion molecule; ICAM-1; inhibitor; leukaemia;  
 KW haematopoietic neoplastic disease; myocardial infarction;  
 KW radiation injury; rheumatoid arthritis; lymphoma metastasis;  
 KW retinoic acid syndrome; all-trans retinoic acid.  
 XX Synthetic.  
 OS  
 XX WO200151508-A1.  
 PN 19-JUL-2001.  
 PD  
 XX 16-JAN-2001; 2001WO-US01382.  
 PF  
 XX 14-JAN-2000; 2000US-0483550.  
 PR  
 XX 16-JAN-2001; 2001US-0760599.  
 PR  
 XX (SCTE-) SCI & TECHNOLOGY CORP @UNM.  
 PA  
 XX Larson RS;  
 PI  
 XX WPI; 2001-432906/46.  
 DR  
 XX Composition comprising a cyclic peptide inhibitor of lymphocyte  
 PT function associated antigen-1 and intracellular adhesion molecule 1  
 PT interaction, for treating e.g. asthma and myocardial infarction -  
 PT  
 XX Example 2; Page 22; 58pp; English.  
 PS  
 XX The sequence represents the amino acid sequence of cyclic peptide  
 XX inhibitor #12 of lymphocyte function associated antigen-1 and  
 CC intracellular adhesion molecule (LFA-1/ICAM-1) interaction. A  
 CC composition comprising a cyclic peptide inhibitor of LFA-1/ICAM-1  
 CC interaction is useful for treating haematopoietic neoplastic disease,  
 CC myocardial infarction, radiation injury, asthma, rheumatoid arthritis or  
 CC lymphoma metastasis. The composition is also useful for inhibiting in a  
 CC subject the interaction between LFA-1 expressed on a leukocyte and ICAM-1  
 CC expressed on another cell, preventing retinoic acid syndrome in a subject  
 CC receiving all-trans retinoic acid, inhibiting growth of leukaemia cells,  
 CC inhibiting emigration of leukocytes from blood into tissue and screening  
 CC a candidate compound for binding to ICAM-1.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 26.2%; Score 32; DB 22; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 4.3e+05;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 15 IHPNSIC 21  
 Db : |||||  
 3 lrpsic 9  
 RESULT 5  
 AAU03751  
 ID AAU03751 standard; peptide; 9 AA.  
 XX  
 AC AAU03751;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE Cyclic peptide inhibitor of LFA-1/ICAM-1 interaction #28.  
 XX  
 KW Cyclic; lymphocyte function associated antigen-1; LFA-1; asthma;  
 KW intracellular adhesion molecule; ICAM-1; inhibitor; leukaemia;  
 KW haematopoietic neoplastic disease; myocardial infarction;  
 KW radiation injury; rheumatoid arthritis; lymphoma metastasis;  
 KW retinoic acid syndrome; all-trans retinoic acid.  
 XX Synthetic.  
 OS

PN WO200151508-A1.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 16-JAN-2001; 2001WO-US01382.  
 XX  
 PR 14-JAN-2000; 2000US-0483550.  
 PR  
 XX 16-JAN-2001; 2001US-0760599.  
 PR  
 XX (SCTE-) SCI & TECHNOLOGY CORP @UNM.  
 PA  
 XX Larson RS;  
 PI  
 XX WPI; 2001-432906/46.  
 DR  
 XX Composition comprising a cyclic peptide inhibitor of lymphocyte  
 PT function associated antigen-1 and intracellular adhesion molecule 1  
 PT interaction, for treating e.g. asthma and myocardial infarction -  
 PT  
 XX Example 2; Page 22; 58pp; English.  
 PS  
 XX The sequence represents the amino acid sequence of cyclic peptide  
 XX inhibitor #28 of lymphocyte function associated antigen-1 and  
 CC intracellular adhesion molecule (LFA-1/ICAM-1) interaction. A  
 CC composition comprising a cyclic peptide inhibitor of LFA-1/ICAM-1  
 CC interaction is useful for treating haematopoietic neoplastic disease,  
 CC myocardial infarction, radiation injury, asthma, rheumatoid arthritis or  
 CC lymphoma metastasis. The composition is also useful for inhibiting in a  
 CC subject the interaction between LFA-1 expressed on a leukocyte and ICAM-1  
 CC expressed on another cell, preventing retinoic acid syndrome in a subject  
 CC receiving all-trans retinoic acid, inhibiting growth of leukaemia cells,  
 CC inhibiting emigration of leukocytes from blood into tissue and screening  
 CC a candidate compound for binding to ICAM-1.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 26.2%; Score 32; DB 22; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 4.3e+05;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 15 IHPNSIC 21  
 Db : |||||  
 3 lrpsic 9  
 RESULT 6  
 AAU36883  
 ID AAU36883 standard; Peptide; 12 AA.  
 XX  
 AC AAU36883;  
 XX  
 DT 11-MAY-1998 (first entry)  
 XX  
 DE HTLV-I protease cleavage recognition site.  
 XX  
 KW Ricin; toxin; antiviral; virucide; retrovirus; protease; HTLV-I;  
 KW cancer; infection; therapy; linker.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9741233-A1.  
 PD  
 XX 06-NOV-1997.  
 XX  
 PF 29-APR-1997; 97WO-CA00288.  
 XX  
 PR 30-APR-1996; 96US-0016509.  
 XX  
 PA (CANG-) CANGENE CORP.  
 XX  
 PI Borgford T;  
 XX

DR WPI; 1997-549735/50.

XX DNAs encoding ricin like toxins A and B - are linked via linker  
 CC containing cleavage site for retroviral protease, used to inhibit or  
 PT destroy mammalian cells infected with retrovirus  
 CC  
 XX  
 PS Claim 5; Page 41; 105pp; English.

XX This claimed peptide is a cleavage recognition site for a HTLV-I  
 CC protease. It is utilised as a linker between the A and B chains  
 CC of a ricin-like protein in a novel recombinant protein. A nucleic  
 CC acid (see AAT97910) encoding such a construct is obtained by PCR  
 CC mutagenesis of the wild-type ricin linker sequence. The invention  
 CC provides novel recombinant proteins which incorporate the A and B  
 CC chains of a ricin-like toxin (preferably the A and B chains of  
 CC ricin) linked by a heterologous linker sequence containing a  
 CC cleavage recognition site for a retroviral protease such as HIV  
 CC protease (see AAW36880-82), HTLV-I (see AAW36883-43) or HTLV-II (see  
 CC AAW36885-65). The recombinant proteins selectively inhibit or  
 CC destroy mammalian cells infected with a retrovirus such as cancer  
 CC cells associated with HTLV or cells associated with HIV. The  
 CC recombinant proteins are non-toxic until the ricin A chain is  
 CC liberated from the B chain by a retroviral protease, and thus can  
 CC be used to specifically target infected cells without the need  
 CC for a cell binding component.

XX Sequence 12 AA;

Query Match 26.2%; Score 32; DB 18; Length 12;

Best Local Similarity 62.5%; Pred. No. 85;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 11 VIPRIHPN 18

Db 5 vipvmhpn 12

RESULT 7

AAW26612  
 ID AAW26612 standard; Peptide; 18 AA.

XX AAW26612;

XX 27-JAN-1998 (first entry)

XX Chick muscle-specific kinase (MuSK) cytoplasmic domain peptide.

XX Receptor tyrosine kinase; muscle specific kinase; MuSK; Dmk;  
 KW chicken; polyclonal antibody.

XX Gallus sp.

XX WO9721811-A2.

XX 19-JUN-1997.

XX 13-DEC-1996; 96WO-US20696.

XX 10-MAY-1996; 96US-0644271.

PR 15-DEC-1995; 95US-0008657.

XX (REGE-) REGENERON PHARM INC.

XX Bowen DC, Glass DJ, Valenzuela DM, Yancopoulos GD;

PI WPI; 1997-332783/30.

XX Nucleotide sequences encoding human agrin and muscle specific kinase  
 PT and related receptor - used in diagnosis and treatment of disorder  
 PT with muscle atrophy

XX Example 10; Page 64; 120pp; English.

XX

CC This peptide sequence comprises the first 18 amino acids of the  
 CC chick muscle specific kinase (MuSK) cytoplasmic domain. It was  
 CC used to raise polyclonal antibodies for MuSK. These were used  
 CC in studies that demonstrated that agrin (see AAW26609) induces  
 CC prominent and rapid tyrosine phosphorylation of MuSK (see also  
 CC AAW26610-11).

XX Sequence 18 AA;

Query Match 26.2%; Score 32; DB 18; Length 18;

Best Local Similarity 30.8%; Pred. No. 1.4e+02;

Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 6 IEGHVVIPRIHPN 18

Db 2 ipseilldrihpn 14

RESULT 8

AAW83548

ID AAW83548 standard; peptide; 9 AA.

XX AAW83548;

XX 04-MAR-1999 (first entry)

XX Tumour necrosis factor receptor N-terminal peptide #6.

XX Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
 KW OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
 KW inflammation; apoptosis.

XX Homo sapiens.

XX WO9849305-A1.

XX 05-NOV-1998.

XX 29-APR-1998; 98WO-US08631.

XX 01-MAY-1997; 97US-0850188.

XX (AMGE-) AMGEN INC.

XX Boyle WJ, Wooden S;

XX WPI; 1999-034661/03.

XX New chimeric osteoprotegerin polypeptides - contain the  
 PT osteoprotegerin dimerisation domain and a heterologous sequence,  
 PT useful to treat TNF and TNFR-mediated disorders

XX Disclosure; Page 16; 92pp; English.

XX The present invention describes a chimeric polypeptide (A1), comprising  
 CC an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
 CC amino acid sequence. Also described are: (1) a multimer polypeptide  
 CC comprising covalently associated A1 monomers; (2) an isolated nucleic  
 CC acid encoding A1; (3) an expression vector comprising the nucleic acid  
 CC sequence; and (4) a host cell transformed or transfected with the  
 CC expression vector so that the nucleic acid is expressible. The products  
 CC from the present invention are useful to treat a variety of disorders  
 CC including those related to receptor binding. Compositions comprising  
 CC tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
 CC are used to treat TNF and TNFR-mediated disorders such as inflammation,  
 CC autoimmune diseases and disorders related to excessive apoptosis. The  
 CC chimeras are also useful for detecting molecules which interact with  
 CC fused heterologous sequences to identify potential new receptors and  
 CC ligands. The present sequence represents a TNFR N-terminal peptide from  
 CC the present invention.

SQ Sequence 9 AA;

Query Match 25.4%; Score 31; DB 20; Length 9;  
Best Local Similarity 77.8%; Pred. No. 4.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 15 IHP--NSIC 21  
||| ||||  
Db 1 ihpqnsic 9

RESULT 9

AAW83549  
ID AAW83549 standard; peptide; 10 AA.

XX  
AC AAW83549;

DT 04-MAR-1999 (first entry)

DE Tumour necrosis factor receptor N-terminal peptide #7.

XX  
KW Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
KW OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
KW inflammation; apoptosis.

XX Homo sapiens.

XX WO9849305-A1.

XX 05-NOV-1998.

XX 29-APR-1998; 98WO-US08631.

XX 01-MAY-1997; 97US-0850188.

XX (AMGE-) AMGEN INC.

XX Boyle WJ, Wooden S;

XX WPI; 1999-034661/03.

XX New chimeric osteoprotegerin polypeptides - contain the  
PT osteoprotegerin dimerisation domain and a heterologous sequence,  
PT useful to treat TNF and TNFR-mediated disorders

XX Disclosure; Page 16; 92pp; English.

XX The present invention describes a chimeric polypeptide (A1), comprising  
CC an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
CC amino acid sequence. Also described are: (1) a multimer polypeptide  
CC comprising covalently associated A1 monomers; (2) an isolated nucleic  
CC acid encoding A1; (3) an expression vector comprising the nucleic acid  
CC sequence; and (4) a host cell transformed or transfected with the  
CC expression vector so that the nucleic acid is expressible. The products  
CC from the present invention are useful to treat a variety of disorders  
CC including those related to receptor binding. Compositions comprising  
CC tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
CC are used to treat TNF and TNFR-mediated disorders such as inflammation,  
CC autoimmune diseases and disorders related to excessive apoptosis. The  
CC chimeras are also useful for detecting molecules which interact with  
CC fused heterologous sequences to identify potential new receptors and  
CC ligands. The present sequence represents a TNFR N-terminal peptide from  
CC the present invention.

XX Sequence 10 AA;

Query Match 25.4%; Score 31; DB 20; Length 10;  
Best Local Similarity 77.8%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 15 IHP--NSIC 21

Db ||| ||||  
2 ihpqnsic 10

RESULT 10

AAW83550  
ID AAW83550 standard; peptide; 11 AA.

XX  
AC AAW83550;

XX 04-MAR-1999 (first entry)

XX Tumour necrosis factor receptor N-terminal peptide #8.

XX  
KW Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
KW OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
KW inflammation; apoptosis.

XX Homo sapiens.

XX WO9849305-A1.

XX 05-NOV-1998.

XX 29-APR-1998; 98WO-US08631.

XX 01-MAY-1997; 97US-0850188.

XX (AMGE-) AMGEN INC.

XX Boyle WJ, Wooden S;

XX WPI; 1999-034661/03.

XX New chimeric osteoprotegerin polypeptides - contain the  
PT osteoprotegerin dimerisation domain and a heterologous sequence,  
PT useful to treat TNF and TNFR-mediated disorders

XX Disclosure; Page 16; 92pp; English.

XX The present invention describes a chimeric polypeptide (A1), comprising  
CC an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
CC amino acid sequence. Also described are: (1) a multimer polypeptide  
CC comprising covalently associated A1 monomers; (2) an isolated nucleic  
CC acid encoding A1; (3) an expression vector comprising the nucleic acid  
CC sequence; and (4) a host cell transformed or transfected with the  
CC expression vector so that the nucleic acid is expressible. The products  
CC from the present invention are useful to treat a variety of disorders  
CC including those related to receptor binding. Compositions comprising  
CC tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
CC are used to treat TNF and TNFR-mediated disorders such as inflammation,  
CC autoimmune diseases and disorders related to excessive apoptosis. The  
CC chimeras are also useful for detecting molecules which interact with  
CC fused heterologous sequences to identify potential new receptors and  
CC ligands. The present sequence represents a TNFR N-terminal peptide from  
CC the present invention.

XX Sequence 11 AA;

Query Match 25.4%; Score 31; DB 20; Length 11;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 15 IHP--NSIC 21  
||| ||||  
Db 3 ihpqnsic 11

RESULT 11

AAW83551  
ID AAW83551 standard; peptide; 12 AA.

XX

AAW83551;  
 04-MAR-1999 (first entry)  
 Tumour necrosis factor receptor N-terminal peptide #9.  
 Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
 OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
 inflammation; apoptosis.  
 Homo sapiens.  
 WO9849305-A1.  
 05-NOV-1998.  
 29-APR-1998; 98WO-US08631.  
 01-MAY-1997; 97US-0850188.  
 (AMGE-) AMGEN INC.  
 Boyle WJ, Wooden S;  
 WPI; 1999-034661/03.  
 New chimeric osteoprotegerin polypeptides - contain the  
 osteoprotegerin dimerisation domain and a heterologous sequence,  
 useful to treat TNF and TNFR-mediated disorders  
 Disclosure; Page 16; 92pp; English.  
 The present invention describes a chimeric polypeptide (A1), comprising  
 an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
 amino acid sequence. Also described are: (1) a multimer polypeptide  
 comprising covalently associated A1 monomers; (2) an isolated nucleic  
 acid encoding A1; (3) an expression vector comprising the nucleic acid  
 sequence; and (4) a host cell transformed or transfected with the  
 expression vector so that the nucleic acid is expressible. The products  
 from the present invention are useful to treat a variety of disorders  
 including those related to receptor binding. Compositions comprising  
 tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
 are used to treat TNF and TNFR-mediated disorders such as inflammation,  
 autoimmune diseases and disorders related to excessive apoptosis. The  
 chimeras are also useful for detecting molecules which interact with  
 fused heterologous sequences to identify potential new receptors and  
 ligands. The present sequence represents a TNFR N-terminal peptide from  
 the present invention.

Query Match 25.4%; Score 31; DB 20; Length 12;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 OY 15 IHP--NSIC 21  
 III IIII  
 Db 4 ihpqnsic 12  
 RESULT 12  
 AAW83552  
 ID AAW83552 standard; peptide; 13 AA.  
 AC AAW83552;  
 XX  
 XX  
 04-MAR-1999 (first entry)  
 Tumour necrosis factor receptor N-terminal peptide #10.  
 Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
 OPG; chimeric; fusion; dimerisation domain; autoimmune disease;

KW inflammation; apoptosis.  
 XX Homo sapiens.  
 XX WO9849305-A1.  
 XX 05-NOV-1998.  
 PD 29-APR-1998; 98WO-US08631.  
 XX 01-MAY-1997; 97US-0850188.  
 PR (AMGE-) AMGEN INC.  
 XX Boyle WJ, Wooden S;  
 XX WPI; 1999-034661/03.  
 XX New chimeric osteoprotegerin polypeptides - contain the  
 osteoprotegerin dimerisation domain and a heterologous sequence,  
 useful to treat TNF and TNFR-mediated disorders  
 Disclosure; Page 16; 92pp; English.  
 The present invention describes a chimeric polypeptide (A1), comprising  
 an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
 amino acid sequence. Also described are: (1) a multimer polypeptide  
 comprising covalently associated A1 monomers; (2) an isolated nucleic  
 acid encoding A1; (3) an expression vector comprising the nucleic acid  
 sequence; and (4) a host cell transformed or transfected with the  
 expression vector so that the nucleic acid is expressible. The products  
 from the present invention are useful to treat a variety of disorders  
 including those related to receptor binding. Compositions comprising  
 tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
 are used to treat TNF and TNFR-mediated disorders such as inflammation,  
 autoimmune diseases and disorders related to excessive apoptosis. The  
 chimeras are also useful for detecting molecules which interact with  
 fused heterologous sequences to identify potential new receptors and  
 ligands. The present sequence represents a TNFR N-terminal peptide from  
 the present invention.

Query Match 25.4%; Score 31; DB 20; Length 13;  
 Best Local Similarity 77.8%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 OY 15 IHP--NSIC 21  
 III IIII  
 Db 5 ihpqnsic 13  
 RESULT 13  
 AAW83553  
 ID AAW83553 standard; peptide; 14 AA.  
 XX AAW83553;  
 AC AAW83553;  
 XX  
 XX 04-MAR-1999 (first entry)  
 Tumour necrosis factor receptor N-terminal peptide #11.  
 Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
 OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
 inflammation; apoptosis.  
 Homo sapiens.  
 WO9849305-A1.  
 05-NOV-1998.



PF 29-APR-1998; 98WO-US08631.  
 XX  
 PR 01-MAY-1997; 97US-0850188.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Boyle WJ, Wooden S;  
 XX  
 DR WPI; 1999-034661/03.  
 XX  
 XX New chimeric osteoprotegerin polypeptides - contain the  
 PT osteoprotegerin dimerisation domain and a heterologous sequence,  
 PT useful to treat TNF and TNFR-mediated disorders  
 XX  
 PS Disclosure; Page 16; 92pp; English.  
 XX  
 CC The present invention describes a chimeric polypeptide (A1), comprising  
 CC an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
 CC amino acid sequence. Also described are: (1) a multimer polypeptide  
 CC comprising covalently associated A1 monomers; (2) an isolated nucleic  
 CC acid encoding A1; (3) an expression vector comprising the nucleic acid  
 CC sequence; and (4) a host cell transformed or transfected with the  
 CC expression vector so that the nucleic acid is expressible. The products  
 CC from the present invention are useful to treat a variety of disorders  
 CC including those related to receptor binding. Compositions comprising  
 CC tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
 CC are used to treat TNF and TNFR-mediated disorders such as inflammation,  
 CC autoimmune diseases and disorders related to excessive apoptosis. The  
 CC chimeras are also useful for detecting molecules which interact with  
 CC fused heterologous sequences to identify potential new receptors and  
 CC ligands. The present sequence represents a TNFR N-terminal peptide from  
 CC the present invention.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 25.4%; Score 31; DB 20; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 QY 15 IHP--NSIC 21  
 DB III IIII  
 6 ihpqnsic 14  
 RESULT 14  
 AAW83554  
 ID AAW83554 standard; peptide; 15 AA.  
 XX  
 AC AAW83554;  
 XX  
 DT 04-MAR-1999 (first entry)  
 XX  
 DE Tumour necrosis factor receptor N-terminal peptide #12.  
 XX  
 KW Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
 KW OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
 KW inflammation; apoptosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9849305-A1.  
 XX  
 PD 05-NOV-1998.  
 XX  
 PF 29-APR-1998; 98WO-US08631.  
 XX  
 PR 01-MAY-1997; 97US-0850188.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Boyle WJ, Wooden S;  
 XX

DR WPI; 1999-034661/03.  
 XX  
 PT New chimeric osteoprotegerin polypeptides - contain the  
 PT osteoprotegerin dimerisation domain and a heterologous sequence,  
 PT useful to treat TNF and TNFR-mediated disorders  
 XX  
 PS Disclosure; Page 16; 92pp; English.  
 XX  
 CC The present invention describes a chimeric polypeptide (A1), comprising  
 CC an osteoprotegerin (OPG) dimerisation domain fused to a heterologous  
 CC amino acid sequence. Also described are: (1) a multimer polypeptide  
 CC comprising covalently associated A1 monomers; (2) an isolated nucleic  
 CC acid encoding A1; (3) an expression vector comprising the nucleic acid  
 CC sequence; and (4) a host cell transformed or transfected with the  
 CC expression vector so that the nucleic acid is expressible. The products  
 CC from the present invention are useful to treat a variety of disorders  
 CC including those related to receptor binding. Compositions comprising  
 CC tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras  
 CC are used to treat TNF and TNFR-mediated disorders such as inflammation,  
 CC autoimmune diseases and disorders related to excessive apoptosis. The  
 CC chimeras are also useful for detecting molecules which interact with  
 CC fused heterologous sequences to identify potential new receptors and  
 CC ligands. The present sequence represents a TNFR N-terminal peptide from  
 CC the present invention.  
 XX  
 SQ Sequence 15 AA;  
 Query Match 25.4%; Score 31; DB 20; Length 15;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 QY 15 IHP--NSIC 21  
 DB III IIII  
 7 ihpqnsic 15  
 RESULT 15  
 AAW35207  
 ID AAW35207 standard; peptide; 16 AA.  
 XX  
 AC AAW35207;  
 XX  
 DT 14-APR-1998 (first entry)  
 XX  
 DE Diastereomer peptide 68.  
 XX  
 KW Diastereomer peptide; infection; therapy; excitatory neurotoxin;  
 KW Honey bee venom; pardaxin; cytolytic activity; cancer;  
 KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;  
 KW agricultural pesticide; cell wall lysis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 3 /note= "D-form residue"  
 FT Misc-difference 5 /note= "D-form residue"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Misc-difference 11 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT Misc-difference 14 /note= "D-form residue"  
 FT Misc-difference 16 /note= "D-form residue"  
 FT

XX WO9731019-A2.  
PN  
XX  
XX 28-AUG-1997.  
PD  
XX  
XX 20-FEB-1997; 97WO-IL000066.  
PF  
XX  
XX 22-FEB-1996; 96IL-0117223.  
PR  
XX  
XX (YEDA ) YEDA RES & DEV CO LTD.  
PA  
XX  
XX Oren Z, Shai Y;  
PI  
XX  
XX WPI; 1997-435088/40.  
DR  
XX  
XX Peptide(s) having selective cytolytic activity - against pathogens  
PT and malignant cells, but no haemolytic activity, used for treating  
PT infections and cancer  
PT  
XX  
XX Example 5; Page 47; 80pp; English.  
PS  
XX  
XX This sequence represents a diastereomer peptide of the  
CC invention. The peptides of the invention have: (a) cytolytic activity on  
CC pathogenic cells (pathogens and malignant cells not naturally present in  
CC the body); but (b) no haemolytic activity, or such activity only at a  
CC concentration significantly higher than that at which they lyse  
CC pathogens. The peptides, their complexes and mixtures are used to treat  
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)  
CC or cancer, in human and veterinary medicine. Also, they can be used as  
CC preservatives for food, cosmetics and agricultural produce, or as  
CC agricultural pesticides. The absence of haemolytic activity (associated  
CC with disturbance of alpha-helical structures) means that the peptides  
CC have few if any toxic effects, and those that include D-aa will have  
CC increased resistance to proteolytic degradation. Non-haemolytic,  
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of  
CC activity, allowing selection of agents for particular applications. Since  
CC these random copolymers induce total lysis of bacterial cell walls,  
CC resistance to them is unlikely to develop.  
XX  
XX Sequence 16 AA;  
SQ

Query Match 25.4%; Score 31; DB 18; Length 16;  
Best Local Similarity 54.5%; Pred. No. 1.7e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 IECHVVIPRIH 16  
|: || | |:  
Db 4 ikahvriirvh 14

Search completed: February 21, 2002, 16:42:59  
Job time: 233 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 16:43:37 ; Search time 31.57 seconds  
(without alignments)  
14.969 Million cell updates/sec

Title: US-08-753-851-14  
Perfect score: 122  
Sequence: 1 CRYGFIQHVIVPRIHPNSIC 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 114342

Minimum DB seq length: 0  
Maximum DB seq length: 21

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	21	2	US-08-143-311B-12
2	104	85.2	19	2	US-08-143-311B-13
3	32	26.2	18	2	US-08-644-271-16
4	31	25.4	19	1	US-08-492-599-3
5	31	25.4	19	5	PCT-US96-10455-3
6	30	24.6	10	3	US-08-159-339A-462
7	30	24.6	13	4	US-09-258-754-101
8	30	24.6	13	4	US-09-042-107-101
9	30	24.6	14	3	US-08-467-023-184
10	30	24.6	14	3	US-08-467-023-186
11	29	23.8	15	1	US-07-854-603-12
12	28.5	23.4	15	4	US-08-743-168B-20
13	28.5	23.4	15	5	PCT-US96-10435-20
14	28.5	23.4	19	2	US-08-374-652C-48
15	28	23.0	11	1	US-07-994-277A-5
16	28	23.0	15	1	US-08-268-251-23
17	28	23.0	15	1	US-07-854-603-14
18	28	23.0	15	5	PCT-US93-01112-23
19	28	23.0	16	2	US-08-310-912A-54
20	28	23.0	16	3	US-08-841-089-54
21	28	23.0	16	4	US-09-301-085-54
22	28	23.0	16	5	PCT-US95-04570-54
23	28	23.0	16	5	PCT-US95-04589-54
24	28	23.0	19	2	US-08-436-420-16
25	28	23.0	20	1	US-07-994-277A-4
26	28	23.0	20	1	US-07-956-700B-12
27	28	23.0	20	1	US-07-956-700B-26

28 23.0 20 1 US-08-476-537-12 Sequence 12, Appl  
29 28 23.0 20 1 US-08-476-537-26 Sequence 26, Appl  
30 28 23.0 20 1 US-08-485-607-12 Sequence 12, Appl  
31 28 23.0 20 1 US-08-485-607-26 Sequence 26, Appl  
32 28 23.0 20 2 US-08-475-879-12 Sequence 12, Appl  
33 28 23.0 20 2 US-08-475-879-26 Sequence 26, Appl  
34 28 23.0 21 1 US-08-732-751-28 Sequence 28, Appl  
35 28 23.0 21 4 US-08-817-548A-3 Sequence 3, Appl  
36 28 23.0 21 6 5283320-4 Patent No. 5283320  
37 27 22.1 8 2 US-08-978-404B-70 Sequence 70, Appl  
38 27 22.1 12 1 US-08-627-497-18 Sequence 18, Appl  
39 27 22.1 12 3 US-08-838-413A-11 Sequence 11, Appl  
40 27 22.1 14 2 US-08-771-602D-33 Sequence 33, Appl  
41 27 22.1 14 4 US-08-896-933-5 Sequence 5, Appl  
42 27 22.1 15 1 US-07-854-603-10 Sequence 10, Appl  
43 27 22.1 17 1 US-08-787-547-27 Sequence 27, Appl  
44 27 22.1 19 1 US-08-442-884-8 Sequence 8, Appl  
45 27 22.1 19 1 US-08-732-751-23 Sequence 23, Appl

## ALIGNMENTS

RESULT 1  
US-08-143-311B-12  
; Sequence 12, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143.311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-12

### RESULT 3

; CITY: York Harbor

/ PATENT NO. 572030  
 / GENERAL INFORMATION:  
 / APPLICANT: Rich, Alexander  
 / APPLICANT: Herbert, Alan  
 / TITLE OF INVENTION: Z-DNA BINDING PROTEIN AND APPLICATIONS  
 / NUMBER OF SEQUENCES: 5  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Kevin M. Farrell, P.C.  
 / STREET: P.O. Box 999  
 / CITY: York Harbor  
 / STATE: ME



QY 6 IEHGVWIPR 14  
: || |:  
Db 2 VNGHQVLPK 10

RESULT 7  
US-09-258-754-101  
: Sequence 101, Application US/09258754  
: Patent No. 6174687  
: GENERAL INFORMATION:  
: APPLICANT: Ruoslahti, Erkki  
: APPLICANT: Pasqualini, Renata  
: APPLICANT: Rajotte, Daniel  
: TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
: TITLE OF INVENTION: Membrane Dipeptidase  
: FILE REFERENCE: P-LJ 3443  
: CURRENT APPLICATION NUMBER: US/09/258,754  
: CURRENT FILING DATE: 1999-02-26  
: EARLIER APPLICATION NUMBER: 09/042,107  
: EARLIER FILING DATE: 1998-03-13  
: NUMBER OF SEQ ID NOS: 452  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 101  
: LENGTH: 13  
: TYPE: PRT  
: ORGANISM: Artificial Sequence  
: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-258-754-101

Query Match 24.6%; Score 30; DB 4; Length 13;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRYGFIEG 8  
: || |:  
Db 1 CRSCGVEG 8

RESULT 8  
US-09-042-107-101  
: Sequence 101, Application US/09042107  
: Patent No. 6232287  
: GENERAL INFORMATION:  
: APPLICANT: Ruoslahti, Erkki  
: APPLICANT: Pasqualini, Renata  
: TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
: TITLE OF INVENTION: Tissues  
: FILE REFERENCE: P-LJ 2892  
: CURRENT APPLICATION NUMBER: US/09/042,107  
: CURRENT FILING DATE: 1998-03-13  
: NUMBER OF SEQ ID NOS: 436  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 101  
: LENGTH: 13  
: TYPE: PRT  
: ORGANISM: Artificial Sequence  
: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-042-107-101

Query Match 24.6%; Score 30; DB 4; Length 13;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRYGFIEG 8  
: || |:  
Db 1 CRSCGVEG 8

RESULT 9  
US-08-467-023-184  
: Sequence 184, Application US/08467023  
: Patent No. 6090386  
: GENERAL INFORMATION:  
: APPLICANT: Griffith, Irwin J.;  
: APPLICANT: Pollock, Joanne;  
: APPLICANT: Bond, Julian F.;  
: APPLICANT: Garman, Richard D;  
: APPLICANT: Kuo, Mei-Chang;  
: APPLICANT: Yeung, Siu-mei H.;  
: APPLICANT: Brauer, Andrew;  
: APPLICANT: Exley, Mark A.;  
: APPLICANT: Powers, Steven P.  
: TITLE OF INVENTION: Allergenic Proteins And Peptides From  
: TITLE OF INVENTION: Japanese Cedar Pollen  
: NUMBER OF SEQUENCES: 261  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
: STREET: 610 Lincoln St  
: CITY: Waltham  
: STATE: MA  
: COUNTRY: USA  
: ZIP: 02154  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/467,023  
: FILING DATE: June 6, 1995  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/350,225  
: FILING DATE: December 6, 1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Jane E. Remillard  
: REGISTRATION NUMBER: 38,872  
: REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (617) 227-7400  
: TELEFAX: (617) 227-5941  
: INFORMATION FOR SEQ ID NO: 184:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 14 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
: FRAGMENT TYPE: internal  
US-08-467-023-184

Query Match 24.6%; Score 30; DB 3; Length 14;  
Best Local Similarity 50.0%; Pred. No. 76;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 GFIEGHVV 11  
: |:  
Db 2 GYFSGHVI 9

RESULT 10  
US-08-467-023-186  
: Sequence 186, Application US/08467023  
: Patent No. 6090386  
: GENERAL INFORMATION:  
: APPLICANT: Griffith, Irwin J.;  
: APPLICANT: Pollock, Joanne;  
: APPLICANT: Bond, Julian F.;  
: APPLICANT: Garman, Richard D;  
: APPLICANT: Kuo, Mei-Chang;  
: APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powers, Steven P.  
TITLE OF INVENTION: Allergenic Proteins And Peptides From  
TITLE OF INVENTION: Japanese Cedar Pollen  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,023  
FILING DATE: June 6, 1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane E. Remillard  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM1-028CPD2)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 186:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-467-023-186

Query Match 24.6%; Score 30; DB 3; Length 14;  
Best Local Similarity 50.0%; Pred. No. 76;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 GFIEGHVV 11  
I: 111;  
Db 2 GYFSGHVI 9

RESULT 11  
US-07-854-603-12  
Sequence 12, Application US/07854603  
Patent No. 5637492  
GENERAL INFORMATION:  
APPLICANT: Dawson, Keith M  
APPLICANT: Edwards, Richard M  
APPLICANT: Forman, Joan M  
TITLE OF INVENTION: Activatable fibrinolytic and  
TITLE OF INVENTION: anti-thrombotic proteins  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. John J. McDonnell  
STREET: Ten South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/854,603  
FILING DATE: 19901207  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: McDonnell, John J  
REGISTRATION NUMBER: 26,949  
REFERENCE/DOCKET NUMBER: 92,338  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-854-603-12

Query Match 23.8%; Score 29; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. NO. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 GFIEGHVV 11  
I: 111;  
Db 5 GYIDGRVV 12

RESULT 12  
US-08-743-168B-20  
Sequence 20, Application US/08743168B  
Patent No. 6271015  
GENERAL INFORMATION:  
APPLICANT: Gilula, No. 6271015ton B  
APPLICANT: Cravatt, Benjamin F  
APPLICANT: Lerner, Richard A  
TITLE OF INVENTION: FATTY-ACID AMIDE HYDROLASE  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute  
STREET: 10550 No. 6271015th Torrey Pines Road  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,168B  
FILING DATE: 04-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/489,535  
FILING DATE: 12-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 485.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal  
US-08-743-168B-20

Query Match 23.4%; Score 28.5; DB 4; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

OY 8 GHVVPRIHPNS 19  
II : II : II :  
Db 5 GHTLIPPL-PNN 15

RESULT 13  
PCT-US96-10435-20  
; Sequence 20. Application PC/TUS9610435  
; GENERAL INFORMATION:  
; APPLICANT: The Scripps Research Institute  
; TITLE OF INVENTION: CIS-9, 10-OCTADECENOAMIDASE  
; NUMBER OF SEQUENCES: 32  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10435  
; FILING DATE: 12-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/489,535  
; FILING DATE: 12-JUN-1995  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
PCT-US96-10435-20

Query Match 23.4%; Score 28.5; DB 5; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

OY 8 GHVVPRIHPNS 19  
II : II : II :  
Db 5 GHTLIPPL-PNN 15

RESULT 14  
US-08-374-652C-48  
; Sequence 48. Application US/08374652C  
; Patent No. 5834286  
; GENERAL INFORMATION:  
; APPLICANT: NEVALAINEN, HELENA K.M.  
; APPLICANT: PALOHEIMO, MARJA T.  
; APPLICANT: FAGERSTROM, RICHARD B.  
; APPLICANT: MIETTINEN-OINONEN, ARJA S.  
; APPLICANT: TURUNEN, MARJA K.  
; APPLICANT: RAMOSEK, JOHN A.  
; APPLICANT: PIDDINGTON, CHRISTOPHER S.  
; APPLICANT: HOUSTON, CHRISTINE S.  
; APPLICANT: CANTRELL, MICHAEL A.  
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,  
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTATE DEGRADING  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON

; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/374,652C  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07058  
; FILING DATE: 27-JUL-1993  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-374-652C-48

Query Match 23.4%; Score 28.5; DB 2; Length 19;  
Best Local Similarity 46.2%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 2 RYGF1-EGHVVIP 13  
III : I : I :  
Db 2 RYGLWNGETVVP 14

RESULT 15  
US-07-994-277A-5  
; Sequence 5. Application US/07994277A  
; Patent No. 5536497  
; GENERAL INFORMATION:  
; APPLICANT: Evans, Richard T.  
; APPLICANT: Bedi, Gurinder S.  
; APPLICANT: Genco, Robert J.  
; APPLICANT: Sojar, Hakimuddin T.  
; TITLE OF INVENTION: Fimbrial Polypeptides  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear  
; STREET: 1800 One M&T Plaza  
; CITY: Buffalo  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 14203-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3 5 inch  
; COMPUTER: IBM compatible/CompuAdd 325TX  
; OPERATING SYSTEM: MS-DOS/Microsoft Windows 3.1  
; SOFTWARE: Wordperfect for Windows 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/994,277A  
; FILING DATE: 19921221  
; CLASSIFICATION: 424



ATTORNEY/AGENT INFORMATION:  
NAME: Nelson, M. Bud  
REGISTRATION NUMBER: 35,300  
REFERENCE/DOCKET NUMBER: 11520.0040  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 856-4000  
TELEFAX: (716) 849-0349  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: peptide  
HYPOTHETICAL: no  
FRAGMENT TYPE: internal  
PUBLICATION INFORMATION:  
RELEVANT RESIDUES IN SEQ ID NO: From 226-236  
RELEVANT RESIDUES IN SEQ ID NO: (without counting the first 10 amino acids which  
RELEVANT RESIDUES IN SEQ ID NO: the leader sequence)  
US-07-994-277A-5

Query Match 23.0%; Score 28; DB 1; Length 11;  
Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 15 IHPNSIC 21  
Db 1 IHPTILC 7

Search completed: February 21, 2002, 16:43:37  
Job time: 171 sec

Query Match 26.9%; Score 28; DB 2; Length 19;  
Best Local Similarity 50.0%; Pred. NO. 2.5e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels

QY 2 YGFIEGHV 9  
|| : ||  
Db 12 YGALEXHI 19

## RESULT 3

B41868

hypothetical protein (traE1 3' region) - Enterococcus faecalis plasmid pADI

C:Species: Enterococcus faecalis

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999

C:Accession: B41868; B37391

R:Pontius, L.F.; Clewell, D.B.

J. Bacteriol. 174, 3152-3160, 1992

A:Title: Conjunctive transfer of Enterococcus faecalis plasmid pADI: nucleotide sequence

A:Reference number: A41868; MUID:92250408

A:Contents: plasmid pADI

A:Accession: B41868

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-15 &lt;PON&gt;

A:Note: sequence extracted from NCBI backbone (NCBIN:99901, NCBI:99906)

R:Clewell, D.B.; Pontius, L.T.; An, F.Y.; Ike, Y.; Suzuki, A.; Nakayama, J.

Plasmid 24, 156-161, 1990

A:Title: Nucleotide sequence of the sex pheromone inhibitor (iAD1) determinant of Enter

A:Reference number: A37391; MUID:91261999

A:Accession: B37391

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-15 &lt;CLUE&gt;

A:Cross-references: GB:M62888; NID:g141853; PIDN:AAA98040.1; PID:g141855

C:Genetics:

A:Genome: plasmid

## Query Match

Best Local Similarity 26.0%; Score 27; DB 2; Length 15;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 HVVIVPR 13

|| |||

Db 5 HVVIVPR 10

## RESULT 4

F57789

gallbladder stone matrix protein 2, 41K - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 03-May-1996

C:Accession: F57789

R:Binette, J.P.; Binette, M.B.

submitted to the Protein Sequence Database, February 1996

A:Description: The proteins of gallbladder stones.

A:Reference number: A57789

A:Accession: F57789

A&gt;Status: preliminary

A:Molecule type: protein

A:Residues: 1-15 &lt;BIN&gt;

A:Note: 9-Phe was also found

## Query Match

Best Local Similarity 26.0%; Score 27; DB 2; Length 15;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFIEG 7

|||||

Db 7 GFIEG 11

## RESULT 5

S12906

protein kinase (EC 2.7.1.37) - starfish (Pisaster ochraceus)

C:Species: Pisaster ochraceus

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Aug-1997

C:Accession: S12904

R:Sanghera, J.S.; Aebersold, R.; Morrison, H.D.; Bures, E.J.; Pelech, S.L.

FEBS Lett. 273, 223-226, 1990

A:Title: Identification of the sites in myelin basic protein that are phosphorylated

A:Reference number: S12904; MUID:91032186

A:Accession: S12904

A&gt;Status: preliminary

A:Molecule type: protein

A:Residues: 1-14 &lt;SAN&gt;

C:Keywords: phosphotransferase

## Query Match

Best Local Similarity 24.0%; Score 25; DB 2; Length 14;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 HVVIVPRIHPNS 18

::|||

Db 1 NLVTPRTTPPS 11

## RESULT 6

A36133

hypothetical protein (bkd 5' region) - Pseudomonas putida (fragment)

C:Species: Pseudomonas putida

C:Date: 30-Nov-1990 #sequence\_revision 30-Nov-1990 #text\_change 30-Sep-1993

C:Accession: A36133

R:Madhusudhan, K.T.; Huang, G.; Burns, G.; Sokatch, J.R.

J. Bacteriol. 172, 5655-5663, 1990

A:Title: Transcriptional analysis of the promoter region of the Pseudomonas putida br

A:Reference number: A36133; MUID:91008935

A:Accession: A36133

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18 &lt;MAD&gt;

A:Cross-references: GB:M33715

## Query Match

Best Local Similarity 24.0%; Score 25; DB 2; Length 18;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFIEG 7

|||

Db 5 GFIEG 9

## RESULT 7

A26999

carboxylesterase (EC 3.1.1.1), intestinal - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 07-Feb-1997

C:Accession: A26999

R:McGhee, J.D.

Biochemistry 26, 4101-4107, 1987

A:Title: Purification and characterization of a carboxylesterase from the intestine o

A:Reference number: A26999; MUID:88000636

A:Accession: A26999

A:Molecule type: protein

A:Residues: 1-13 &lt;MCG&gt;

C:Keywords: carboxylic ester hydrolase; intestine

## Query Match

Best Local Similarity 23.1%; Score 24; DB 2; Length 13;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YGFIEG 7

|| : ||

Db 8 YGKVEG 13

RESULT 8  
A47393  
neuropeptide callatostatin 1 - bluebottle fly (Calliphora vomitoria)  
N:Contains: neuropeptide callatostatin 2  
C:Species: Calliphora vomitoria  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47393; B47393; C47393  
R:Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993  
A:Title: Callatostatin: neuropeptides from the blowfly Calliphora vomitoria with sequen  
A:Reference number: A47393; MUID:93211980  
A:Accession: A47393  
A:Molecule type: protein  
A:Residues: 1-16 <DUV>  
A:Experimental source: thoracic ganglia, brains, heads  
A:Note: sequence extracted from NCBI backbone (NCBIP:128476)  
A:Accession: B47393  
A:Molecule type: protein  
A:Residues: 3-16 <DU2>  
A:Experimental source: head  
A:Note: sequence extracted from NCBI backbone (NCBIP:128477)  
A:Accession: C47393  
A:Molecule type: protein  
A:Residues: 1-8 <DU3>  
A:Experimental source: thoracic ganglia  
A:Note: sequence extracted from NCBI backbone (NCBIP:128478)  
C:Function:  
A:Description: inhibits juvenile hormone production  
C:Keywords: amidated carboxyl end; neuropeptide  
F:1-16/Product: callatostatin 1 #status experimental <MA1>  
F:3-16/Product: callatostatin 2 #status experimental <MA2>  
F:9-16/Product: callatostatin 3 #status experimental <MA3>  
F:16/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 23.1%; Score 24; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYGF 4  
| | | |  
DB 11 RYGF 14

RESULT 9  
DRDPPP  
distal-retinal-pigment hormone - northern shrimp  
N:Alternate names: light-adapting hormone  
C:Species: Pandalus borealis (northern shrimp)  
C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 23-Aug-1996  
C:Accession: A01473  
R:Fernlund, P.  
Biochim. Biophys. Acta 439, 17-25, 1976  
A:Title: Structure of a light-adapting hormone from the shrimp, Pandalus borealis.  
A:Reference number: A01473; MUID:76253762  
A:Accession: A01473  
A:Molecule type: protein  
A:Residues: 1-18 <FER>  
C:Comment: This peptide causes migration of the distal retinal pigment into the proximal  
the amount of light entering the retinulas.  
C:Superfamily: pigment-dispersing hormone  
C:Keywords: amidated carboxyl end; hormone  
F:18/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 23.1%; Score 24; DB 1; Length 18;  
Best Local Similarity 41.7%; Pred. No. 1.1e+03;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 GFIEGHVPIRI 14  
| | | | |  
DB 3 GMINSLGIPRV 14

RESULT 10  
C30503  
Ig gamma-2b chain C region (F5.5.1) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 16-Aug-1996  
C:Accession: C30503  
R:Gillmore, G.L.; Bard, J.A.; Birshstein, B.K.  
J. Immunol. 141, 1754-1761, 1988  
A:Title: DNA rearrangements affecting both variable and constant regions of Ig H chain  
A:Reference number: A30503; MUID:88315788  
A:Accession: C30503  
A:Molecule type: mRNA  
A:Residues: 1-12 <GIL>  
A:Experimental source: myeloma cell line  
C:Keywords: immunoglobulin

Query Match 22.1%; Score 23; DB 2; Length 12;  
Best Local Similarity 50.0%; Pred. No. 1e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 GHVVIPRI 14  
| | | | |  
DB 5 GHRTSPRL 12

RESULT 11  
PN0149  
beta-Gliadine 13 - Aegilops longissima (fragment)  
C:Species: Aegilops longissima  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: PN0149  
R:Odintsova, T.I.; Egorov, T.A.  
Biokhimiia 55, 509-516, 1990  
A:Title: N-Terminal sequences of omega-gliadins of Aegilops longissima: On the origin  
A:Reference number: PN0146; MUID:90283493  
A:Accession: PN0149  
A:Molecule type: protein  
A:Residues: 1-18 <ODI>  
A:Experimental source: strain K-202  
C:Superfamily: gliadin

Query Match 22.1%; Score 23; DB 2; Length 18;  
Best Local Similarity 30.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 VVIPRIHPNS 18  
| | | | |  
DB 3 VPVPQLQPQN 12

RESULT 12  
I40062  
shikimate 5-dehydrogenase (EC 1.1.1.25) - Buchnera aphidicola (fragment)  
C:Species: Buchnera aphidicola  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 08-Oct-1999  
C:Accession: I40062  
R:Roubbakhsh, D.; Baumann, P.  
Gene 155, 107-112, 1995  
A:Title: Characterization of a putative 23S-5S rRNA operon of Buchnera aphidicola (en  
A:Reference number: I40061; MUID:95212914  
A:Accession: I40062  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-18 <RES>  
A:Cross-references: EMBL:U10496; NID:9854711; PIDN:AAA79125.1; PID:9854712  
C:Genetics:  
A:Gene: aroE  
C:Keywords: oxidoreductase

Query Match 22.1%; Score 23; DB 2; Length 18;  
Best Local Similarity 45.5%; Pred. No. 1.6e+03;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 FIEGHVVPRI 14  
| | | | |  
DB 2 FFLWHNVLPPI 12

## RESULT 13

B61218  
alpha-gliadin 6Ha - grass (Haynaldia villosa) (fragment)  
C:Species: Haynaldia villosa, Dasyphyrum villosum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Dec-1997 #text\_change 17-Mar-1999  
C:Accession: B61218  
R:Shewry, P.R.; Sabelli, P.A.; Parmar, S.; Lafandra, D.  
Biochem. Genet. 29, 207-211, 1991  
A:Title: alpha-type prolamins are encoded by genes on chromosomes 4Ha and 6Ha of Haynaldia villosa  
A:Reference number: A61218; MUID:91315394  
A:Accession: B61218  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SHE>  
C:Keywords: seed; storage protein

Query Match 21.2%; Score 22; DB 2; Length 10;  
Best Local Similarity 37.5%; Pred. No. 1.3e+03;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VVIPRIHP 16  
| | | | |  
DB 3 VVPQLQP 10

## RESULT 14

PA0037  
plastocyanin 2 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 23-Mar-1995  
C:Accession: PA0037  
R:Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
submitted to JIPID, July 1994  
A:Description: Separation and characterization of Arabidopsis proteins by two-dimensional gel electrophoresis  
A:Reference number: PA0001  
A:Accession: PA0037  
A:Molecule type: protein  
A:Residues: 1-12 <KAM>  
A:Experimental source: stem

Query Match 21.2%; Score 22; DB 2; Length 12;  
Best Local Similarity 66.7%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 GHVVIP 12  
| | | | |  
DB 2 GXVLP 7

## RESULT 15

PT0304  
Ig heavy chain CRD3 region (clone 5-115B) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0304  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J<sub>H</sub> gene segments in the development of the B cell repertoire  
A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0304  
A:Molecule type: DNA

A:Residues: 1-13 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 21.2%; Score 22; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYGFIE 6  
| | | | |  
DB 1 RGYVD 6

Search completed: February 21, 2002, 17:16:49  
Job time: 109 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:16:35 ; Search time 10.06 Seconds  
(without alignments)  
69.248 Million cell updates/sec

Title: US-08-753-851-15

Perfect score: 104

Sequence: 1 RYGFIEGHVVIPRIHPNSI 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 968

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	23.1	16	1 ALL1_CALVO	P41839 calliphora
2	24	23.1	18	1 DRPH_PANBO	P01209 pandalus bo
3	23	22.1	19	1 UKAI1_HUMAN	P31940 homo sapien
4	22	21.2	12	1 PKV2_PERAM	P81555 periplaneta
5	22	21.2	16	1 ARCD_PSEPU	P41147 pseudomonas
6	22	21.2	18	1 ARGF_STRRM	P80575 streptomyce
7	21.5	20.7	15	1 RM12_YEAST	P36522 saccharomyc
8	21	20.2	10	1 ANGT_CHICK	P01018 gallus gall
9	21	20.2	10	1 BPP8_BOTIN	P30426 bothrops in
10	21	20.2	10	1 GON3_ONCKE	P20367 oncorhynch
11	21	20.2	11	1 TNK3_PSEGU	P42988 pseudophryn
12	21	20.2	13	1 LMT4_LOCMI	P41490 locusta mig
13	20	19.2	9	1 NEUX_HUMAN	P04277 homo sapien
14	20	19.2	15	1 UC08_WAIZE	P80614 zea mays (m
15	20	19.2	19	1 FIBB_PIG	P14477 sus scrofa
16	19	18.3	8	1 ANG2_BOTJA	Q10582 bothrops ja
17	19	18.3	9	1 NEUX_RAT	P11382 rattus norv
18	19	18.3	10	1 GLEM_HUMAN	P02728 homo sapien
19	19	18.3	14	1 ANGT_HORSE	P01016 equus cabal
20	19	18.3	19	1 CAT3_FASHE	P80532 fasciola he
21	19	18.3	19	1 LPRM_STAAT	P30363 staphylococ
22	18	17.3	9	1 PKK1_PERAM	P82691 periplaneta
23	18	17.3	9	1 TAL3_PICJA	P17441 pichia jadi
24	18	17.3	10	1 AH3_PRUSE	P29261 prunus sero
25	18	17.3	10	1 ANGT_BOTJA	Q10581 bothrops ja
26	18	17.3	10	1 ANGT_BOVIN	P01017 bos taurus
27	18	17.3	11	1 ANGT_CRIGE	P09037 crinia geor
28	18	17.3	11	1 BPPB_AKHA	P01021 agkistrodon
29	18	17.3	11	1 MGS_THETS	P41989 theromyzon
30	18	17.3	11	1 TINI_HOPTI	P82651 hoplobatrac
31	18	17.3	13	1 AH4_PRUSE	P29262 prunus sero
32	18	17.3	13	1 TP13_PHYRO	P04096 phyllomedus
33	18	17.3	14	1 COCO_LIMPO	P35586 limulus pol

34 18 17.3 14 1 DCMM\_PSECF P19914 pseudomonas  
35 18 17.3 14 1 HV14\_PIG P01155 sus scrofa  
36 18 17.3 15 1 DCMM\_PSECH P19917 pseudomonas  
37 18 17.3 15 1 BEIA\_MICCR P81266 microplitis  
38 18 17.3 16 1 AHL\_PRUSE P29259 prunus sero  
39 18 17.3 16 1 FOR1\_MYRGU P81438 myrmecia gu  
40 18 17.3 17 1 BOLA\_MEGPE P07495 megabombus  
41 18 17.3 18 1 FIXA\_RHILE P14313 rhizobium l  
42 18 17.3 18 1 TOPL\_KLEAE P46155 klebsiella  
43 18 17.3 19 1 FRHG\_METBA P80491 methanosarc  
44 17 16.3 9 1 LMT3\_LOCMI P41489 locusta mig  
45 17 16.3 10 1 BPP2\_BOTIN P30422 bothrops in

#### ALIGNMENTS

RESULT 1  
ALL1\_CALVO STANDARD; PRT; 16 AA.  
AC P41839;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CALLATOSTATIN 1 (LEU-CALLATOSTATIN 1) [CONTAINS: CALLATOSTATIN 2 (LEU-CALLATOSTATIN 2); CALLATOSTATIN 3 (LEU-CALLATOSTATIN 3)].  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
SEQUENCE.  
RP TISSUE=Thoracic ganglion, Brain, and Head;  
RX MEDLINE=93211980; PubMed=8460157;  
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S., Thorpe A.;  
RA "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with sequence homology to cockroach allatostatins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
RN [2]  
CHARACTERIZATION.  
RP MEDLINE=94291167; PubMed=8020069;  
RA Duve H., Thorpe A.;  
RA "Distribution and functional significance of Leu-callatostatins in the blowfly Calliphora vomitoria.";  
RL Cell Tissue Res. 276:367-379(1994).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.  
CC -!- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEBRAL COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY SYSTEM AND INTESTINE.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
DR PIR; A47393; A47393.  
DR PIR; B47393; B47393.  
DR PIR; C47393; C47393.  
KW Neuropeptide; Amidation.  
FT PEPTIDE 1 16 CALLATOSTATIN 1.  
FT PEPTIDE 3 16 CALLATOSTATIN 2.  
FT PEPTIDE 9 16 CALLATOSTATIN 3.  
FT MOD\_RES 16 16 AMIDATION.  
SQ SEQUENCE 16 AA; 1907 MW; A435B68C26EC3D09 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 16;

Best Local Similarity 100.0%; Pred. No. 6.5e+02; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0;

QY 1 RYGF 4  
||||

Db 11 RYGF 14

## RESULT 2

DRPH\_PANBO STANDARD; PRT; 18 AA.  
 AC P01209;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE PIGMENT-DISPERSING HORMONE (PDH) (LIGHT ADAPTING DISTAL RETINAL  
 DE PIGMENT HORMONE) (DRPH).  
 OS Pandalus borealis (Northern red shrimp).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidae;  
 OC Pandalidae; Pandalus.  
 OX NCBI\_TaxID=6703;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76253762; PubMed=952951;  
 RA Fernlund P.;  
 RT "Structure of a light-adapting hormone from the shrimp, Pandalus  
 RT borealis";  
 RL Biochim. Biophys. Acta 439:17-25(1976).  
 CC -!- FUNCTION: CAUSES THE MIGRATION OF THE DISTAL RETINAL PIGMENT  
 CC INTO THE PROXIMAL END OF THE PIGMENT CHROMATOPHORE CELLS AND  
 CC THUS DECREASES THE AMOUNT OF LIGHT ENTERING THE RETINULAS.  
 CC -!- SIMILARITY: TO THE PDH OF OTHER ARTHROPODS.  
 DR PIR; A01473; DRDPPT.  
 KW Hormone; Amidation.  
 FT MOD\_RES 18 18 AMIDATION.  
 SQ SEQUENCE 18 AA; 1903 MW; B2346B6D0178650E CRC64;

Query Match 23.1%; Score 24; DB 1; Length 18;

Best Local Similarity 41.7%; Pred. No. 7.3e+02;

Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 GFIEGHVVIPRI 14

I I I I I I I

Db 3 GMINSILGIPRV 14

## RESULT 3

UKAL\_HUMAN STANDARD; PRT; 19 AA.  
 AC P31940;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF EPIDERMAL KERATINOCYTES (SPOT 1118)  
 DE (FRAGMENTS).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Keratinocytes;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Cellis J.E.,  
 RA Vandekerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RT protein database of normal human epidermal keratinocytes";  
 RL Electrophoresis 13:960-969(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 7.24, ITS MW IS: 23.5 KDA.  
 DR Aarhus/Ghent-2DPAGE; 1118; IEF.  
 FT NON\_TER 1 1  
 FT UNSURE 6 6  
 FT NON\_CONS 6 7  
 FT NON\_CONS 12 13  
 FT UNSURE 17 17

FT NON\_TER 19  
SQ SEQUENCE 19 AA; 2087 MW; EF7515F79D50DE12 CRC64;Query Match 22.1%; Score 23; DB 1; Length 19;  
Best Local Similarity 25.0%; Pred. No. 1.1e+03;  
Matches 3; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 8 HWVIPRIHPNSI 19

I I I I I I I

Db 1 HIGLVRLTPTEV 12

## RESULT 4

PVK2\_PERAM STANDARD; PRT; 12 AA.  
 AC P81555;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PERIVISCEROKININ-2 (PEA-PVK-2).  
 OS Periplaneta americana (American cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blattellidae; Blattidae; Periplaneta.  
 OX NCBI\_TaxID=6978;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE=Abdominal perisymphathetic organs;  
 RX MEDLINE=98326577; PubMed=9663444;  
 RA Predel R., Rapus J., Eckert M., Holman G.M., Nachman R.J., Wang Y.,  
 RA Penzlin H.;  
 RT "Isolation of periviscerokinin-2 from the abdominal perisymphathetic  
 RT organs of the American cockroach, Periplaneta americana";  
 RL Peptides 19:801-809(1998).  
 CC -!- FUNCTION: MYOACTIVE PEPTIDE; HAS EXCITATORY ACTIONS ON THE  
 CC HYPERNEURAL MUSCLE.  
 CC -!- MASS SPECTROMETRY: MW=1189.3; METHOD=MALDI.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 12 12 AMIDATION.  
 SQ SEQUENCE 12 AA; 1190 MW; 2F4D8EEE1EB05728 CRC64;

Query Match 21.2%; Score 22; DB 1; Length 12;

Best Local Similarity 33.3%; Pred. No. 1e+03;  
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 GFIEGHVVIPRI 14

I I I I I I I

Db 1 GSSSGLISMPRV 12

## RESULT 5

ARCD\_PSEPU STANDARD; PRT; 16 AA.  
 AC P41147;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ARGININE/ORNITHINE ANTIPOPTER (FRAGMENT).  
 GN ARCD.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 4359;  
 RA Wilson S.D., Wang M., Filpula D.;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGININE  
 CC AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE  
 CC ARGININE DEIMINASE PATHWAY.



```

CC CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC CC (POTENTIAL).
CC CC -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC CC PERMEASES.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U07185; AAA16963.1;
CC CC Transport; Amino-acid transport; Transmembrane; Inner membrane.
CC CC NON_TER 1
CC CC SEQUENCE 16 AA; 1644 MW; 90B48A7C8FAA9705 CRC64;

Query Match 21.2%; Score 22; DB 1; Length 16;
Best Local Similarity 30.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YGFIEGHVVI 11
   || : : :
DB 7 YGLYDGFLL 16

RESULT 6
ID AROF_STRM STANDARD; PRT; 18 AA.
AC P80575;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PHOSPHO-3-DEHYDRO-3-DEOXYHEPTONATE ALDOLASE (EC 4.1.2.15) (PHOSPHO-2-
DE KETO-3-DEOXYHEPTONATE ALDOLASE) (DAHP SYNTHETASE) (3-DEOXY-D-ARABINO-
DE HEPTULOSONATE 7-PHOSPHATE SYNTHASE) (FRAGMENT).
GN AROH.
OS Streptomyces rimosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1927;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RC STRAIN-ISOLATE 4018;
RX MEDLINE=96349103; PubMed=8760910;
RA Walker G.E., Dunbar B., Hunter I.S., Nimmo H.G., Coggins J.R.;
RT "Evidence for a novel class of microbial
RT 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase in Streptomyces
RT coelicolor A3(2), Streptomyces rimosus and Neurospora crassa.";
RL Microbiology 142:1973-1982(1996).
CC -!- CATALYTIC ACTIVITY: 7-PHOSPHO-2-DEHYDRO-3-DEOXY-D-ARABINO-
CC HEPTONATE + ORTHOPHOSPHATE = PHOSPHOENOLPYRUVATE + D-BRITHROSE
CC 4-PHOSPHATE + H(2)O.
CC -!- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: BELONGS TO CLASS-II DAHP SYNTHETASE FAMILY.
DR InterPro: IPR002480; DAHP_synth_2.
DR Pfam: PF01474; DAHP_synth_2; 1.
CC Aromatic amino acid biosynthesis; Lyase.
FT NON_TER 1
FT NON_TER 18
CC SEQUENCE 18 AA; 2209 MW; 7FE627E713C871BB CRC64;

Query Match 21.2%; Score 22; DB 1; Length 18;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFTEGH 8

```

```

DB 13 GFEEVH 18
   || | |
RESULT 7
ID RM12_YEAST STANDARD; PRT; 15 AA.
AC P36522;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L12 (YML12) (FRAGMENT).
GN MRPL12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
RX MEDLINE=91285106; PubMed=2060626;
RA Grohmann L., Graack H.-R., Kruft V., Choli T., Goldschmidt-Reisin S.,
RA Kitakawa M.;
RT "Extended N-terminal sequencing of proteins of the large ribosomal
RT subunit from yeast mitochondria.";
RL FEBS Lett. 284:51-56(1991).
DR PIR; S17261; S17261.
DR SGD; L0002687; MRPL12.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 15
FT NON_TER 15
CC SEQUENCE 15 AA; 1851 MW; 74BCD9FEDDD3900 CRC64;

Query Match 20.7%; Score 21.5; DB 1; Length 15;
Best Local Similarity 46.2%; Pred. No. 1.6e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 4 FIEGHVVIPIRHP 16
   || | | | |
DB 6 FVE---VIVRFNP 15

RESULT 8
ID ANGT_CHICK STANDARD; PRT; 10 AA.
AC P01018;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ANGIOTENSINOGEN [CONTAINS: ANGIOTENSIN I; ANGIOTENSIN II] (FRAGMENT).
GN SERPIN A8 OR AGT.
OS Gallus gallus (Chicken), and
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 93934;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken;
RX MEDLINE=74127845; PubMed=4361802;
RA Nakayama T., Nakajima T., Sokabe H.;
RT "Comparative studies on angiotensins. 3. Structure of fowl
RT angiotensin and its identification by DNS-method.";
RL Chem. Pharm. Bull. 21:2085-2087(1973).
RN [2]
RP SEQUENCE.
RC SPECIES=C.japonica;
RX MEDLINE=90284684; PubMed=2191893;
RA Takei Y., Hasegawa Y.;
RT "Vasopressor and depressor effects of native angiotensins and
RT inhibition of these effects in the Japanese quail.";
RL Gen. Comp. Endocrinol. 79:12-22(1990).
CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN

```

```

CC CLEAVES ANGIOTENSIN I. FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC BALANCE OF BODY FLUIDS.
CC -1- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR PIR; A01250; A01250.
DR PIR; A0917; A0917.
DR PIR; A0624; A0624.
DR InterPro: IPR000215; Serpin.
DR PROSITE: PS00284; SERPIN; PARTIAL.
KW Vasoconstrictor; Plasma; Serpin.
FT PEPTIDE 1 10 ANGIOTENSIN I.
FT PEPTIDE 1 8 ANGIOTENSIN II.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1232 MW; CEFBEDD761F2DB42 CRC64;

Query Match 20.2%; Score 21; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 IHPNSI 19
Db 5 VHPFSL 10

RESULT 9
BPP8_BOTIN STANDARD; PRT; 10 AA.
AC P30426;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S5,1 (ANGIOTENSIN-CONVERTING
DE ENZYME INHIBITOR).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE-Venom.
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; H37196; H37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 20.2%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 HPN 17
Db 5 HPN 7

RESULT 10
GON3_ONCKE STANDARD; PRT; 10 AA.
ID GON3_ONCKE
AC P20367;P81751;

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GONADOLIBERIN III (GONADOTROPIN-RELEASING HORMONE III) (GNRH-III) (LH-
DE RH III) (LULIBERIN III).
OS GnRH3.
GN Oncorhynchus keta (Chum salmon), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O. keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C. pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC PIR; A21114; A21114.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
FT SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 20.2%; Score 21; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YGFIEG 7
Db 5 YGWLPG 10

RESULT 11
TKN3_PSEGU STANDARD; PRT; 11 AA.
ID TKN3_PSEGU
AC P42988;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE KASSININ-LIKE PEPTIDE K-III (PG-KIII).
OS Pseudophryne guentheri (Frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC Pseudophryne.
OX NCBI_TaxID=30349;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=90287814; PubMed=2356157;
RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA Roberts J.D., Melchiorri P., Erspamer V.;
RT "Six novel tachykinin- and bombesin-related peptides from the skin of
RT the Australian frog Pseudophryne guentheri.";
RL Peptides 11:299-304(1990).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,

```

CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR: D60409; D60409.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Amidation.  
 KW MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 11 11 AMIDATION.  
 FT SEQUENCE 11 AA; 1268 MW; 3DBA7C37C9CB1457 CRC64;  
 SQ

Query Match 20.2%; Score 21; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 HPN 17  
 DB 3 HPN 5

RESULT 12  
 LMT4\_LOCOMI STANDARD; PRT; 13 AA.  
 ID LMT4\_LOCOMI  
 AC P41490;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE LOCUSTAMYTROPIN 4 (LOM-MT-4).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Brain;  
 RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,  
 RA de Loof A.;  
 RT "Isolation, identification and synthesis of locustamytropin III and  
 RT IV, two additional neuropeptides of Locusta migratoria: members of the  
 RT locustamytropin peptide family.";  
 RL Insect Biochem. Mol. Biol. 22:447-452(1992).  
 CC -!- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTILE ACTIVITY  
 CC (MYOTROPIC ACTIVITY). LOM-MT IV SEEMS TO BE A MORE POTENT MUSCLE  
 CC STIMULATOR THAN LOM-MT I, II AND III.  
 CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 DR InterPro: IPR001484; Pyrokinin.  
 DR PROSITE: PS00539; PYROKININ; 1.  
 DR Neuropeptide; Amidation; Pyrokinin.  
 KW MOD\_RES 13 13 AMIDATION.  
 FT MOD\_RES 13 AA; 1553 MW; 20861943824D6698 CRC64;  
 SQ

Query Match 20.2%; Score 21; DB 1; Length 13;  
 Best Local Similarity 42.9%; Pred. No. 1.6e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 13 RIHPNSI 19  
 DB 1 RLHONG 7

RESULT 13  
 NEUX\_HUMAN STANDARD; PRT; 9 AA.  
 ID NEUX\_HUMAN  
 AC P04277;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 20-MAR-1987 (Rel. 04, Last sequence update)

DT 01-JUL-1989 (Rel. 11, Last annotation update)  
 DE NEUROPENSIN-RELATED PEPTIDE (NRP) (KINETENSIN).  
 OS Homo sapiens (Human); Bos taurus (Bovine), and  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606, 9913, 9986;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=Human;  
 RX MEDLINE=86242180; PubMed=3087352;  
 RA Mogard M.H., Kobayashi R., Chen C.F., Lee T.D., Reeve J.R. Jr.,  
 RA Shively J.E., Walsh J.H.;  
 RT "The amino acid sequence of kinetensin, a novel peptide isolated from  
 RT pepsin-treated human plasma: homology with human serum albumin,  
 RT neurotensin and angiotensin.";  
 RL Biochem. Biophys. Res. Commun. 136:983-988(1986).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=Human, Bovine, and Rabbit;  
 RX MEDLINE=87194805; PubMed=2437111;  
 RA Carraway R.E., Mitra S.P., Cochran D.E.;  
 RT "Structure of a biologically active neurotensin-related peptide  
 RT obtained from pepsin-treated albumin(s).";  
 RL J. Biol. Chem. 262:5968-5973(1987).  
 CC -!- FUNCTION: REGULATION OF FAT DIGESTION, LIPID ABSORPTION, AND  
 CC BLOOD FLOW (POTENTIAL).  
 CC -!- SIMILARITY: STRUCTURALLY WITH NEUROPENSIN AND ANGIOTENSIN I.  
 CC STRONG SEQUENCE HOMOLOGY WITH RAT NRP.  
 DR PIR: A03239; ABHUSK.  
 DR PIR: A26693; A26693.  
 KW Hormone.  
 SQ SEQUENCE 9 AA; 1172 MW; C804DB4761F4140D CRC64;

Query Match 19.2%; Score 20; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 IPRHP 16  
 DB 1 IARRHP 6

RESULT 14  
 UC08\_MAIZE STANDARD; PRT; 15 AA.  
 ID UC08\_MAIZE  
 AC P80614;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLOGICAL COLEOPTILE (SPOT 159)  
 DE (FRAGMENT).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Coleoptile;  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Pernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated  
 RT genome analysis program.";  
 RL Theor. Appl. Genet. 93:997-1005(1996).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.4, ITS MW IS: 38.8 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE  
 CC FAMILY.  
 DR Maize-2DPAGE; P80614; COLEOPTILE.  
 DR MaizeDB: 123934; -.  
 FT NON\_TER 1 1

FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1785 MW; 1978BD6AB4DDF8D CRC64;

Query Match 19.2%; Score 20; DB 1; Length 15;  
Best Local Similarity 42.9%; Pred. No. 2.7e+03;  
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YGFIEGH 8  
| : ||  
Db 8 YPVVPGH 14

## RESULT 15

FIBB\_PIG  
ID FIBB\_PIG STANDARD; PRT; 19 AA.  
AC P14477;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-JAN-1990 (Rel. 13, Last annotation update)  
DE FIBRINOPEPTIDE B.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE.  
RA Blomback B., Blomback M., Grondahl N.J.; "  
RT "Studies on fibrinopeptides from mammals."  
RL Acta Chem. Scand. 19:1789-1791(1965).  
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT..  
DR InterPro: IPR002181; Fibrinogen\_C.  
DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT MOD\_RES 4 4 SULFATION.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2200 MW; 09F87E44F4F3863D CRC64;

Query Match 19.2%; Score 20; DB 1; Length 19;  
Best Local Similarity 50.0%; Pred. No. 3.4e+03;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 PRIH 15  
| : |  
Db 12 PKVH 15

Search completed: February 21, 2002, 17:19:31  
Job time: 176 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:16:15 ; Search time 22.04 Seconds  
(without alignments)  
126.097 Million cell updates/sec

Title: US-08-753-851-15  
Perfect score: 104  
Sequence: 1 RYGFIEGHVVIPRIHPNSI 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 4358

Minimum DB seq length: 0  
Maximum DB seq length: 19

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mmc.\*  
8: sp\_organalle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rident.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	63	60.6	12	4 Q9UC29	Q9uc29 homo sapien
2	31	29.8	11	5 Q9TWM2	Q9twm2 aplysia cal
3	30	28.8	16	2 Q44610	Q44610 buchnera ap
4	29	27.9	17	4 Q9UC28	Q9uc28 homo sapien
5	28	26.9	18	4 Q9UCF4	Q9ucf4 homo sapien
6	28	26.9	19	4 Q9UC82	Q9uc82 homo sapien
7	27	26.0	15	2 Q52135	Q52135 enterococcu
8	26	25.0	9	8 P92072	P92072 euhadra her
9	26	25.0	19	5 P82600	P82600 aedes aegyp
10	25	24.0	17	4 Q9UCN0	Q9ucn0 homo sapien
11	24	23.1	10	13 Q9PRY8	Q9pry8 tritakis scy
12	24	23.1	17	2 Q9R4C4	Q9r4c4 agrobacteri
13	24	23.1	17	4 Q95795	Q95795 homo sapien
14	24	23.1	17	7 P79480	P79480 cervus elap
15	24	23.1	17	13 P82396	P82396 litorea ran
16	23	22.1	14	6 Q9061	Q9061 bos taurus
17	22	21.2	10	4 Q90CQ4	Q9ucq4 homo sapien
18	22	21.2	10	12 Q75595	Q75595 human immun
19	22	21.2	11	6 Q9XSP8	Q9xsp8 presbytis j

20	22	21.2	11	6 Q9XSP7	Q9xsp7 pygathrix n
21	22	21.2	11	6 Q9XSP6	Q9xsp6 pongo pygma
22	22	21.2	11	6 Q9XSP2	Q9xsp2 hylobates s
23	22	21.2	13	4 Q9NR93	Q9nr93 homo sapien
24	22	21.2	16	2 Q99374	Q99374 staphylococ
25	22	21.2	17	4 Q9UCA4	Q9uca4 homo sapien
26	22	21.2	17	7 Q19716	Q19716 homo sapien
27	22	21.2	17	10 P82671	P82671 spinacia ol
28	22	21.2	19	2 Q07479	Q07479 streptococc
29	22	21.2	19	4 Q9BVX6	Q9bvx6 homo sapien
30	22	21.2	19	5 Q9TWH4	Q9twh4 heliothis v
31	22	21.2	19	6 Q9BD06	Q9bdd6 ovis aries
32	22	21.2	19	11 Q9QV20	Q9qv20 mus sp. per
33	22	21.2	19	13 Q42416	Q42416 gallus gall
34	21.5	20.7	14	12 Q86578	Q86578 sigma virus
35	21	20.2	11	6 Q9XSQ4	Q9xsq4 gorilla gor
36	21	20.2	11	6 Q9XSP5	Q9xsp5 pan troglod
37	21	20.2	14	2 Q9R5P6	Q9r5p6 legionella
38	21	20.2	15	6 Q9TR14	Q9tr14 bos taurus
39	21	20.2	17	12 Q85719	Q85719 reovirus sp
40	21	20.2	18	10 Q39829	Q39829 glycine max
41	21	20.2	19	2 Q9X6X7	Q9x6x7 salmonella
42	21	20.2	19	4 Q16271	Q16271 homo sapien
43	21	20.2	19	4 Q9UCE4	Q9uce4 homo sapien
44	21	20.2	19	10 Q82685	Q82685 gerbera hyb
45	20.5	19.7	9	12 Q67605	Q67605 squash leaf

## ALIGNMENTS

RESULT 1

Q9UC29 ID Q9UC29 PRELIMINARY; PRT; 12 AA.  
AC Q9UC29;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE 85 KDA GLIOMA MEMBRANE PROTEIN/CD44 HOMOLOG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RX MEDLINE=95079448; Pubmed=7527301;  
RA Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;  
RT "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes  
overexpressed CD44 in glioma cells.";  
RL Cancer Immunol. Immunother. 39:313-317(1994).  
SQ SEQUENCE 12 AA; 1337 MW; 2E0F6CE9D9D2C1E8 CRC64;

Query Match 60.6%; Score 63; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFIEGHVVIPRI 14

DB 1 GFIEGHVVIPRI 12

RESULT 2

Q9TWM2 ID Q9TWM2 PRELIMINARY; PRT; 11 AA.  
AC Q9TWM2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE 85 KDA GLIOMA MEMBRANE PROTEIN/CD44 HOMOLOG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RX MEDLINE=95079448; Pubmed=7527301;  
RA Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;  
RT "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes  
overexpressed CD44 in glioma cells.";  
RL Cancer Immunol. Immunother. 39:313-317(1994).  
SQ SEQUENCE 12 AA; 1337 MW; 2E0F6CE9D9D2C1E8 CRC64;

```

OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE.
RX MEDLINE=95083478; PubMed=7991459;
RA Villim F.S., Cropper E.C., Rosen S.C., Tenenbaum R., Kupfermann I.,
RA Weiss K.R.;
RT "Structure, localization, and action of buccalin B: a bioactive
RT peptide from Aplysia.";
RL Peptides 15:959-969(1994).
SQ SEQUENCE 11 AA; 1153 MW; 692253F9C9C86B44 CRC64;

Query Match 29.8%; Score 31; DB 5; Length 11;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYGFIEG 7
Db 4 RYGFVGG 10

RESULT 3
Q44610 PRELIMINARY; PRT; 16 AA.
ID Q44610;
AC Q44610;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE SHIKIMATE DEHYDROGENASE (FRAGMENT).
GN AROE.
OS Buchnera aphidicola.
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=9;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95212914; PubMed=7535281;
RA Roubhakhsh D., Baumann P.;
RT "Characterization of a putative 23S-5S rRNA operon of Buchnera
RT aphidicola (endosymbiont of aphids) unlinked to the 16S rRNA-encoding
RT gene.";
RL Gene 155:107-112(1995).
DR EMBL; U10499; AAA79128.1; -.
FT NON_TER 1
SQ SEQUENCE 16 AA; 1891 MW; 72A8175598D30DF1 CRC64;

Query Match 28.8%; Score 30; DB 2; Length 16;
Best Local Similarity 38.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 FIEGHVVIPRIHP 16
Db 2 FLEFHGIFPRIEP 14

RESULT 4
Q9UC28 PRELIMINARY; PRT; 17 AA.
ID Q9UC28;
AC Q9UC28;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE 85 KDA GLIOMA MEMBRANE PROTEIN/CD44 HOMOLOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=95079448; PubMed=7527301;
RA Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;
RT "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes
RT overexpressed CD44 in glioma cells.";
```

```

RL Cancer Immunol. Immunother. 39:313-317(1994).
SQ SEQUENCE 17 AA; 1788 MW; 7806F51C97A3E217 CRC64;

Query Match 27.9%; Score 29; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 HPNSI 19
Db 1 HPNSI 5

RESULT 5
Q9UCF4 PRELIMINARY; PRT; 18 AA.
ID Q9UCF4;
AC Q9UCF4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CHYMOTRYPSINOGEN HOMOLOG (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93285747; PubMed=8509158;
RA Kim Y.J., Zhou Z., Hurtado J., Wood D.L., Choi A.S., Pescovitz M.D.,
RA Warfel K.A., Vandagriff J., Davis J.K., Kwon B.S.;
RT "IDM patients' sera recognize a novel 30-kD pancreatic autoantigen
RT related to chymotrypsinogen."
RL Immunol. Invest. 22:219-227(1993).
SQ SEQUENCE 18 AA; 1826 MW; 38105D1037CA6837 CRC64;

Query Match 26.9%; Score 28; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 7.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IPR1HP 16
Db 3 VPA1HP 8

RESULT 6
Q9UC82 PRELIMINARY; PRT; 19 AA.
ID Q9UC82;
AC Q9UC82;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 22 KDA STONE MATRIX PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=95215817; PubMed=7701298;
RA Binette J.P., Binette M.B.;
RT "Sequencing of proteins extracted from stones.";
RL Scanning Microsc. 8:233-239(1994).
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSPSP; P17670; LIDS.
DR InterPro; IPR001189; SOD_M1.
DR Pfam; PF00081; sodfe; 1.
KW Oxidoreductase.
SQ SEQUENCE 19 AA; 2162 MW; AD703A074F3C9655 CRC64;
```

Query Match	25.0%	Score 26;	DB 8;	Length 9;
Best Local Similarity	33.3%	Pred. No. 4.7e+05;		

RESULT	10
Q9UCN0	
ID	Q9UCN0
AC	PRELIMINARY;
PRT	17 AA.
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE	TRANSFERIN RECEPTOR.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata;
OC	Mammalia; Eutheria; Primates; Carnarhini; Hominidae;
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE.

RX MEDLINE=92375195; PubMed=1380674;  
 RA Chicx R.M., Urban R.G., Lane W.S., Gorga J.C., Stern L.J.,  
 RA Vignali D.A., Strominger J.L.;  
 RT "Predominant naturally processed peptides bound to HLA-DR1 are derived  
 from MHC-related molecules and are heterogeneous in size.";  
 RL Nature 358:764-768(1992).  
 SQ SEQUENCE 17 AA; 2035 MW; A7DEDA39A2538A88 CRC64;

Query Match 24.0%; Score 25; DB 4; Length 17;  
 Best Local Similarity 33.3%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 IEHVVIPRIHP 16  
 :|:|:|  
 Db 2 VEYHFLSPYVSP 13

RESULT 11  
 Q9PRY8 PRELIMINARY; PRT; 10 AA.  
 AC Q9PRY8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE ANGIOTENSIN I.  
 OS Triakis scyllium (leopard shark) (Triakis scyllia).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;  
 OC Triakis.  
 OX NCBI\_TaxID=30494;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94141412; PubMed=8308464;  
 RA Takel Y., Hasegawa Y., Watanabe T.X., Nakajima K., Hazon N.;  
 RT "A novel angiotensin I isolated from an elasmobranch fish.";  
 RL J. Endocrinol. 139:281-285(1993).  
 SQ SEQUENCE 10 AA; 1284 MW; 20F02FD761E04B47 CRC64;

Query Match 23.1%; Score 24; DB 13; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 PRIHP 16  
 :|:|:|  
 Db 3 PYIHP 7

RESULT 12  
 Q9R4C4 PRELIMINARY; PRT; 17 AA.  
 AC Q9R4C4;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE PROCATHECHUATE 3,4-DIOXYGENASE TYPE II BETA SUBUNIT (EC 1.13.11.3)  
 DE (FRAGMENT).  
 OS Agrobacterium radiobacter.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=358;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=963337865; PubMed=8772173;  
 RA Hammer A., Stolz A., Knackmuss H.;  
 RT "Purification and characterization of a novel type of protocatechuate  
 3,4-dioxygenase with the ability to oxidize 4-sulfocatechol.";  
 RL Arch. Microbiol. 166:92-100(1996).  
 SQ SEQUENCE 17 AA; 1812 MW; ECA688148553EDF2 CRC64;

Query Match 23.1%; Score 24; DB 2; Length 17;

Best Local Similarity 28.6%; Pred. No. 3.2e+03;  
 Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 FIEGHWIPRIHPN 17  
 :|:|:|:|  
 Db 3 FLPGXPEVPAEYPS 16

RESULT 13  
 Q95795 PRELIMINARY; PRT; 17 AA.  
 AC Q95795;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE ADVANCED GLYCOSYLATION END PRODUCT-SPECIFIC RECEPTOR (FRAGMENT).  
 GN RAGE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Blazkova M., Kankova K.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF065211; AAD15888.1;  
 KW Receptor.  
 FT NON\_TER 1 1 S -> G.  
 FT VARIANT 9 9  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1727 MW; 38C7EE8959E00D72 CRC64;

Query Match 23.1%; Score 24; DB 4; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 3.2e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 IPRIHPNS 18  
 :|:|:|:|  
 Db 2 VARVLPSN 9

RESULT 14  
 P79480 PRELIMINARY; PRT; 17 AA.  
 ID P79480;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
 DE MHC CLASS II DRB (FRAGMENT).  
 OS Cervus elaphus (Red deer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervioidea;  
 OC Cervidae; Cervinae; Cervus.  
 OX NCBI\_TaxID=9860;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Swarbrick P.A., Crawford A.M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63083; AAB37783.1;  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 2022 MW; 2B7EBBCF17F004FA CRC64;

Query Match 23.1%; Score 24; DB 7; Length 17;  
 Best Local Similarity 55.6%; Pred. No. 3.2e+03;  
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 YGFIEGHVV 10  
 :|:|:|:|  
 Db 5 YGVIESFTV 13



```

RESULT 15
P82396
ID P82396 PRELIMINARY; PRT; 17 AA.
AC P82396;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE AUREIN 3.3/3.3.1.
OS Litoria raniformis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=116057;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=GRANULAR DORSAL GLAND;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT Australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2."
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST L.LACTIS, M.LUTEUS,
CC PP.MULTOCIDA, S.AUREUS, S.EPIDERMIDIS AND S.UBERIS. PROBABLY ACTS
CC BY DISTURBING MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC SHOWS ANTICANCER ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT PEPTIDE 3 16 AUREIN 3.3.1.
FT MOD_RES 17 17 AMIDATION.
SQ SEQUENCE 17 AA; 1797 MW; F6AC0A471428EAFD CRC64;

```

```

Query Match      23.1%; Score 24; DB 13; Length 17;
Best Local Similarity 66.7%; Pred. NO. 3.2e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 IEHVV 10
Db 9 IAGHIV 14

```

Search completed: February 21, 2002, 17:19:15  
Job time: 180 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:13:00 ; Search time 23.85 Seconds  
(without alignments)  
59.010 Million cell updates/sec

Title: US-08-753-851-15

Perfect score: 104  
Sequence: 1 RYGFIEGHVVIPRIHPNSI 19

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 193517

Minimum DB seq length: 0  
Maximum DB seq length: 19

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
  - 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
  - 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
  - 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
  - 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
  - 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
  - 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
  - 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
  - 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
  - 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
  - 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
  - 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
  - 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
  - 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
  - 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
  - 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
  - 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
  - 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
  - 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
  - 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
  - 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
  - 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	19	15	AA53485
2	32	30.8	12	18	AAW36883
3	32	30.8	18	18	AAW26612
4	31	29.8	16	18	AAW35207
5	31	29.8	16	19	AAW82923
6	31	29.8	16	21	AAW17458
7	30	28.8	10	15	AA138164
8	30	28.8	10	20	AA45732
9	30	28.8	10	21	AA23536
10	30	28.8	14	15	AA53696
11	30	28.8	15	17	AA97942

12	30	28.8	15	17	AA97943	Japan cedar pollen
13	30	28.8	15	21	AA57842	Human leptin recep
14	30	28.8	18	22	AA19282	Peptide #5716 enco
15	30	28.8	18	22	AA32006	Peptide #6043 enco
16	29	27.9	15	10	AA90252	Antigenic peptide
17	28.5	27.4	15	18	AA23520	Purified cis-9,10-
18	28.5	27.4	15	19	AA57799	Fatty acid amide h
19	28.5	27.4	15	22	AA47117	EPHF containing pe
20	28	26.9	12	22	AA59728	Tomato ACC synthas
21	28	26.9	14	22	AA98109	Human SNP associat
22	28	26.9	14	22	AA88174	CD66 peptide CD66f
23	28	26.9	15	14	AA39535	Calmodulin inhibit
24	28	26.9	15	21	AA98888	HLA class II bindi
25	28	26.9	15	22	AA84543	MAGE2 DR supermoti
26	28	26.9	15	22	AA84862	MAGE2 DR supermoti
27	28	26.9	18	19	AA69517	Protein SEQ ID NO:
28	28	26.9	19	22	AA21662	Peptide #8096 enco
29	28	26.9	19	22	AA37969	Peptide #12006 enc
30	27	26.0	8	22	AA22804	HIV peptide SEQ ID
31	27	26.0	8	22	AA22805	HIV peptide SEQ ID
32	27	26.0	8	22	AA23396	HIV peptide SEQ ID
33	27	26.0	12	16	AA83348	KB-binding random
34	27	26.0	12	17	AA49225	Human leucocyte an
35	27	26.0	12	18	AAW36886	HTLV-II protease c
36	27	26.0	12	21	AA93879	Reactive peptide w
37	27	26.0	13	21	AA15711	Porphyromonas ging
38	27	26.0	14	16	AAW21490	Rhesus apolipoprot
39	27	26.0	15	19	AAW77212	Pharmaceutically a
40	27	26.0	17	19	AAW78817	HBS presl protein
41	27	26.0	17	21	AA33628	MHC class II assoc
42	27	26.0	17	22	AA93724	HBS presl peptide
43	27	26.0	18	19	AA45582	Peptide fragment o
44	27	26.0	19	14	AA31569	HTLV-II epitope, H
45	27	26.0	19	20	AA26945	IS3/RP, a D35E mot

ALIGNMENTS

RESULT 1	
AA53485	
ID	AA53485 standard; peptide; 19 AA.
XX	AA53485;
AC	AA53485;
XX	01-DEC-1994 (first entry)
DT	XX
XX	XX
DE	CD44 peptide CD44-12.
XX	XX
KW	Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;
KW	arthritis; tumor cell metastasis; autoimmune disease;
KW	immunosuppressive.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO9409811-A.
XX	XX
PD	11-MAY-1994.
XX	XX
PF	29-OCT-1993; 93WO-US10412.
XX	XX
PR	30-OCT-1992; 92US-0973339.
XX	XX
PA	(UYDU-) UNIV DUKE.
XX	XX
PI	Hale LP, Haynes BF, Liao H, Patton KL, Telen MU;
XX	XX
DR	WPI; 1994-167121/20.
XX	XX
PT	Use of CD44 protein and new peptide derivs - for developing prods
PT	for inflammation, immune-mediated tissue damage and tumour cell
PT	metastasis
XX	XX

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 76-96 of the CD44 protein sequence.

XX SQ Sequence 19 AA;

Query Match 100.0%; Score 104; DB 15; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.2e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYGFIEGHVVIPIHPNSI 19  
| | | | | | | | | | | | | | | | | | |  
Db 1 rygfieghvviipihpnsi 19

RESULT 2  
AAW36883  
ID AAW36883 standard; Peptide; 12 AA.  
XX  
AC AAW36883;

XX 11-MAY-1998 (first entry)

XX HTLV-I protease cleavage recognition site.

XX Ricin; toxin; antiviral; virucide; retrovirus; protease; HTLV-I; cancer; infection; therapy; linker.

XX Synthetic.

XX W09741233-A1.

XX 06-NOV-1997.

XX 29-APR-1997; 97WO-CA00288.

XX 30-APR-1996; 96US-0016509.

XX (CANG-) CANGENE CORP.

XX Borgford T;

XX WPI; 1997-549735/50.

XX DNAs encoding ricin like toxins A and B - are linked via linker containing cleavage site for retroviral protease, used to inhibit or destroy mammalian cells infected with retrovirus

XX Claim 5; Page 41; 105pp; English.

XX This claimed peptide is a cleavage recognition site for a HTLV-I protease. It is utilised as a linker between the A and B chains of a ricin-like protein in a novel recombinant protein. A nucleic acid (see AAT97910) encoding such a construct is obtained by PCR mutagenesis of the wild-type ricin linker sequence. The invention provides novel recombinant proteins which incorporate the A and B chains of a ricin-like toxin (preferably the A and B chains of ricin) linked by a heterologous linker sequence containing a cleavage recognition site for a retroviral protease such as HIV protease (see AAW36880-82), HTLV-I (see AAW36883-43) or HTLV-II (see AAW36885-65). The recombinant proteins selectively inhibit or destroy mammalian cells infected with a retrovirus such as cancer cells associated with HTLV or cells associated with HIV. The recombinant proteins are non-toxic until the ricin A chain is liberated from the B chain by a retroviral protease, and thus can be used to specifically target infected cells without the need for a cell binding component.

XX SQ Sequence 12 AA;

Query Match 30.8%; Score 32; DB 18; Length 12;  
Best Local Similarity 62.5%; Pred. No. 48;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 VIPRIHPN 17  
| | | | |  
Db 5 vlpvmbpnm 12

RESULT 3  
AAW26612  
ID AAW26612 standard; Peptide; 18 AA.  
XX  
AC AAW26612;

XX 27-JAN-1998 (first entry)

XX Chick muscle-specific kinase (MuSK) cytoplasmic domain peptide.  
XX Receptor tyrosine kinase; muscle specific kinase; MuSK; Dmk;  
KW chicken; polyclonal antibody.

XX Gallus sp.

XX W09721811-A2.

XX 19-JUN-1997.

XX 13-DEC-1996; 96WO-US20696.

XX 10-MAY-1996; 96US-0644271.

XX 15-DEC-1995; 95US-0008657.

XX (REGE-) REGENERON PHARM INC.

XX Bowen DC, Glass DJ, Valenzuela DM, Yancopoulos GD;

XX WPI; 1997-332783/30.

XX Nucleotide sequences encoding human agrin and muscle specific kinase and related receptor - used in diagnosis and treatment of disorder with muscle atrophy

XX Example 10; Page 64; 120pp; English.

XX This peptide sequence comprises the first 18 amino acids of the chick muscle specific kinase (MuSK) cytoplasmic domain. It was used to raise polyclonal antibodies for MuSK. These were used in studies that demonstrated that agrin (see AAW26609) induces prominent and rapid tyrosine phosphorylation of MuSK (see also AAW26610-11).

XX SQ Sequence 18 AA;

Query Match 30.8%; Score 32; DB 18; Length 18;  
Best Local Similarity 30.8%; Pred. No. 78;  
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 IEGHVVIPIHPN 17  
| | | | |  
Db 2 lpsellldrlhpn 14

RESULT 4  
AAW35207  
ID AAW35207 standard; peptide; 16 AA.  
XX  
AC AAW35207;

XX 14-APR-1998 (first entry)  
 XX Diastereomer peptide 68.  
 XX Diastereomer peptide; infection; therapy; excitatory neurotoxin;  
 KW Honey bee venom; pardaxin; cytolytic activity; cancer;  
 KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;  
 KW agricultural pesticide; cell wall lysis.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 3 /note= "D-form residue"  
 FT Misc-difference 5 /note= "D-form residue"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Misc-difference 11 /note= "D-form residue"  
 FT Misc-difference 13 /note= "D-form residue"  
 FT Misc-difference 14 /note= "D-form residue"  
 FT Misc-difference 16 /note= "D-form residue"  
 XX WO9731019-A2.  
 XX 28-AUG-1997. 97WO-IL00066.  
 XX 20-FEB-1997; 97WO-IL00066.  
 XX 22-FEB-1996; 96IL-0117223.  
 XX (YEDA ) YEDA RES & DEV CO LTD.  
 XX Oren Z, Shai Y;  
 XX WPI; 1997-435088/40.  
 XX Peptide(s) having selective cytolytic activity - against pathogens  
 XX and malignant cells, but no haemolytic activity, used for treating  
 XX infections and cancer  
 XX Example 5; Page 47; 80pp; English.  
 XX This sequence represents a diastereomer peptide of the  
 XX invention. The peptides of the invention have: (a) cytolytic activity on  
 XX pathogenic cells (pathogens and malignant cells not naturally present in  
 XX the body); but (b) no haemolytic activity, or such activity only at a  
 XX concentration significantly higher than that at which they lyse  
 XX pathogens. The peptides, their complexes and mixtures are used to treat  
 XX infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)  
 XX or cancer, in human and veterinary medicine. Also, they can be used as  
 XX preservatives for food, cosmetics and agricultural produce, or as  
 XX agricultural pesticides. The absence of haemolytic activity (associated  
 XX with disturbance of alpha-helical structures) means that the peptides  
 XX have few if any toxic effects, and those that include D-aa will have  
 XX increased resistance to proteolytic degradation. Non-haemolytic,  
 XX cytotoxic random copolymers of pardaxin, each has a specific spectrum of  
 XX activity, allowing selection of agents for particular applications. Since  
 XX these random copolymers induce total lysis of bacterial cell walls,  
 XX resistance to them is unlikely to develop.  
 XX Sequence 16 AA;

Query Match 29.8%; Score 31; DB 18; Length 16;  
 Best Local Similarity 54.5%; Pred. No. 1e+02;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 5 IEGHVVIPRIH 15  
 Db 4 ikahvrihrv 14  
 I: | | | | |  
 | : | | | | |  
 RESULT 5  
 AAW82923  
 ID AAW82923 standard; peptide; 16 AA.  
 XX AC AAW82923;  
 XX DT 19-MAY-1999 (first entry)  
 XX DE Antipathogenic peptide.  
 XX KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;  
 KW cancer; infection; disinfectant; contact lens wetting solution;  
 KW preservative; pesticide; fungicide; bactericide.  
 XX OS Synthetic.  
 XX FN WO9837090-A1.  
 XX PD 27-AUG-1998.  
 XX PF 19-FEB-1998; 98WO-IL00081.  
 XX PR 20-FEB-1997; 97WO-IL00066.  
 XX PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX PI Oren Z, Shai Y;  
 XX DR WPI; 1998-594464/50.  
 XX PT New non-haemolytic cytolytic agent useful in treating cancer or  
 XX infections - is a peptide comprising a moiety which disrupts the  
 XX continuity of an alpha-helical structure  
 XX Example 5; Page 46; 126pp; English.  
 XX The present peptide is used to produce the agents of the invention. The  
 XX specification describes a non-haemolytic, cytolytic agent, which is a  
 XX peptide, a complex of bundled peptides, a mixture of peptides or a random  
 XX peptide copolymer. The agent has a selective cytolytic activity on  
 XX pathogenic cells. The agent is selected from a cyclic derivative of a  
 XX peptide which has a net positive charge greater than 1, comprises L-amino  
 XX acid residues and/or D-amino acid residues and comprises an alpha-helix  
 XX breaker moiety, or a peptide (or cyclic derivative of this) which  
 XX comprises L-amino acid residues and D-amino acid residues, has a net  
 XX positive charge greater than 1 and has an amino acid sequence such that  
 XX a corresponding amino acid sequence comprising only L-amino acid residues  
 XX is not found in nature. The cytolytic agents may be used for treatment of  
 XX cancer or for treatment of several diseases caused by pathogens,  
 XX including bacterial, fungal, viral, mycoplasma and protozoan infections.  
 XX They may be used in both human and veterinary medicine. They may also be  
 XX used as disinfectants for destruction of microorganisms, i.e. in  
 XX solutions for wetting contact lenses, as preservatives, e.g., in the  
 XX cosmetic and food industries, as pesticides (e.g. fungicides or  
 XX bactericides) or for preservation of agricultural products.  
 XX Sequence 16 AA;  
 SQ  
 Query Match 29.8%; Score 31; DB 19; Length 16;  
 Best Local Similarity 54.5%; Pred. No. 1e+02;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 5 IEGHVVIPRIH 15

Db 4 ikahvriirvh 14

RESULT 6  
AAB17458  
ID AAB17458 standard; Peptide: 16 AA.

XX AC AAB17458;  
XX DT 31-OCT-2000 (first entry)  
XX DE Antipathogenic peptide sequence SEQ ID NO:562.

XX KW Modified peptide: therapeutic agent; fusion; Fc domain; cancer;  
XX KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
XX KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
XX KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
XX KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
XX KW vascular endothelial growth factor; matrix metalloproteinase;  
XX KW asthma; thrombosis; pharmaceutical.

XX OS Synthetic.

XX PN WO200024782-A2.

XX PD 04-MAY-2000.

XX PF 25-OCT-1999; 99WO-US25044.

XX PR 23-OCT-1998; 98US-0105371.

XX PR 22-OCT-1999; 99US-0428082.

XX PA (AMGE-) AMGEN INC.

XX PI Feige U, Liu C, Cheetham J, Boone TC;

XX PS WPI; 2000-350702/30.

XX PT Novel composition of matter comprising an Fc domain and  
XX PT pharmacologically active peptides, useful for treating cancer and  
XX PT autoimmune diseases -

XX PS Claim 39; Page 392; 608pp; English.

XX CC The present invention describes composition of matter (I) comprising an  
XX CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
XX CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
XX CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,  
XX CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4  
XX CC where P1, P2, P3, and P4 = are each independently sequences of  
XX CC pharmacologically active peptides; L1, L2, L3, and L4 = are each  
XX CC independently linkers; and a, b, c, d, e, and f = are each independently  
XX CC 0 or 1, provided that at least 1 of a and b is 1. The composition can  
XX CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive  
XX CC activities. DNAs, vectors and host cells from the present invention can  
XX CC be used for producing pharmaceutical compositions. The compositions are  
XX CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
XX CC The use of an Fc domain (rather than a Fab domain) can provide a longer  
XX CC half-life or incorporate functions such as Fc receptor binding, protein  
XX CC A binding, complement fixation, and possibly placental transfer. AAA69443  
XX CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid  
XX CC sequences used in the exemplification of the present invention.

XX SQ Sequence 16 AA;

Query Match 29.88; Score 31; DB 21; Length 16;  
Best Local Similarity 54.5%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 IEGHVIPRIH 15  
1: ||| |:

Db 4 ikahvriirvh 14

RESULT 7  
AAY38164  
ID AAY38164 standard; Peptide: 10 AA.

XX AC AAY38164;  
XX DT 29-SEP-1999 (first entry)  
XX DE Hepatitis B virus-derived HLA-binding peptide.

XX KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;  
XX KW MHC; major histocompatibility complex; viral infection; anticancer;  
XX KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX OS Hepatitis B virus.

XX PN WO9403205-A1.

XX PD 17-FEB-1994.

XX PF 06-AUG-1993; 93WO-US07421.

XX PR 05-MAR-1993; 93US-0027746.

XX PR 07-AUG-1992; 92US-0926666.

XX PA (CYTE-) CYTEL CORP.

XX PI Cellis E, Grey HM, Kubo RT, Sette A;

XX PS WPI; 1994-065403/08.

XX PT Peptide which specifically binds selected MHC allele - used to  
XX PT induce an immune response for treatment or prevention of viral  
XX PT infection or cancer, or for diagnosis

XX PS Disclosure; Page 107; 150pp; English.

XX CC The sequence is a specific example of a group of new immunogenic  
XX CC peptides having an HLA-A3.2, HLA-A1, or HLA-A24.1 binding  
XX CC motif. For example, the peptides having an HLA-A3.2 binding motif  
XX CC each have 9-10 residues and contain, from the N-terminus to the  
XX CC C-terminus, (a) a first conserved residue selected from L, M, I,  
XX CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of  
XX CC K, R, Y, H or F, where the first and second conserved residues are  
XX CC separated by 6-7 residues. The peptides are capable of binding  
XX CC selected MHC molecules and inducing an immune response. They can be  
XX CC used to treat and/or prevent viral infection and cancer, e.g. prostate  
XX CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce  
XX CC antibodies for use as diagnostic or therapeutic agents. The peptides  
XX CC can also be used as diagnostic agents.

XX SQ Sequence 10 AA;

Query Match 28.8%; Score 30; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 83;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 IEGHVIPR 13  
: ||| |:

Db 2 vngqvlpk 10

RESULT 8  
AAY45732  
ID AAY45732 standard; Peptide: 10 AA.

XX AC AAY45732;

XX DT 01-DEC-1999 (first entry)

XX Immunogenic peptide having a human leukocyte antigen binding motif #343.  
DE Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
KW immune response; T cell activation; major histocompatibility complex;  
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
KW vaccine; immunisation.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
XX WO9945954-A1.  
PN  
XX 16-SEP-1999.  
PD  
XX 13-MAR-1998; 98WO-US05039.  
PF  
XX 13-MAR-1998; 98WO-US05039.  
PR  
XX (EPIM-) EPIMUNE INC.  
PA  
XX Sette A, Kubo RT, Sidney J, Cellis E, Grey HM, Southwood S;  
PI WPI; 1999-551214/46.  
XX  
XX New immunogenic peptides with HLA binding motif, useful in treatment  
PT and diagnosis of cancers and viral diseases -  
PT  
XX Claim 1; Page 40; 150pp; English.  
XX  
CC AAY45390 to AAY49214 represent specifically claimed immunogenic peptides  
CC having a human major histocompatibility complex (MHC) Class I (also  
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
CC response against the antigen from which the peptide is derived.  
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
CC normally induced by an antigen in the form of a peptide fragment bound  
CC to a HLA molecule, rather than the intact foreign antigen itself, and  
CC are particularly important in tumour rejection and in fighting viral  
CC infections. The peptides are therefore useful therapeutically to treat  
CC or prevent viral infections and cancers in mammals (especially humans)  
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
CC They can be administered as vaccines to elicit an immune response in  
CC individuals susceptible or otherwise at risk of viral infection or  
CC cancer, or used to treat chronic or acute conditions. They are also  
CC useful diagnostically, and can be used to induce a cytotoxic T cell  
CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
CC produce CTLs ex vivo for infusion back into a patient. The  
CC polynucleotides encoding the immunogenic peptides are also useful  
CC therapeutically and for immunisation as above.  
XX  
XX Sequence 10 AA;  
SQ

Query Match 28.8%; Score 30; DB 20; Length 10;  
Best Local Similarity 44.4%; Pred. No. 83;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 IEGHVIVPR 13  
: ||| |:  
Db 2 vnghqvlpk 10

RESULT 9  
AAB23536  
ID AAB23536 standard; peptide; 10 AA.  
AC AAB23536;  
XX  
XX 02-JAN-2001 (first entry)  
DT  
XX

DE Angiogenic vasculature homing peptide SEQ ID 13.  
XX  
KW Angiogenic vasculature homing peptide; proteoglycan; NG2; tumour;  
KW drug targeting; cytotoxic; chemotherapeutic agent.  
XX  
OS Unidentified.  
XX  
XX WO200048464-A1.  
PN  
XX 24-AUG-2000.  
PD  
XX 16-FEB-2000; 2000WO-US03864.  
PF  
XX 16-FEB-1999; 99US-0250700.  
PR  
XX (BURN-) BURNHAM INST.  
PA  
XX Burg MA, Pasqualini R, Arap W, Ruoslahti EI, Stallcup WB;  
PI WPI; 2000-558237/51.  
XX  
XX New angiogenic vasculature homing peptides and conjugates e.g. for  
PT targeting drugs to tumors -  
PT  
XX Example 1; Page 58; 86pp; English.  
XX  
CC Sequences AAB23524-B23539 represent angiogenic vasculature homing  
CC peptides, which interact with proteoglycan NG2. The invention includes a  
CC method for targeting angiogenic vasculature in a tumour in vivo, and a  
CC method for inhibiting angiogenesis in a tumour comprising administering  
CC one of the angiogenic vasculature homing peptides linked to a moiety such  
CC as a drug or cytotoxic agent. The peptides exhibit cytostatic activity,  
CC and are used for targeting drugs, cytotoxic agents or chemotherapeutic  
CC agents to tumours or tissues or other organs associated with angiogenic  
CC vasculature.  
XX  
XX Sequence 10 AA;  
SQ

Query Match 28.8%; Score 30; DB 21; Length 10;  
Best Local Similarity 50.0%; Pred. No. 83;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 EGHVVIVPR 13  
: ||| |:  
Db 1 dahvlvpr 8

RESULT 10  
AAR53696  
ID AAR53696 standard; peptide; 14 AA.  
XX  
XX AAR53696;  
AC  
XX 01-FEB-1995 (first entry)  
DT  
XX Japanese cedar pollen allergen Cry j IIB fragment.  
DE  
XX Cedar pollinosis; diagnostic.  
KW  
XX Synthetic.  
OS  
XX WO9411512-A.  
PN  
XX 26-MAY-1994.  
PD  
XX 12-NOV-1993; 93WO-US11000.  
PF  
XX 12-NOV-1992; 92US-0975179.  
PR  
XX (IMMU-) IMMULOGIC PHARM CORP.  
PA  
XX Brauer A, Kuo M, Pollock J, Yeung S;  
PI

XX WPI; 1994-183513/22.  
 XX Allergenic Cry j II protein and fragments from Japanese cedar  
 PT pollen - used to diagnose, treat and prevent Japanese cedar  
 PT pollinosis  
 XX  
 PS Disclosure; Page 38; 89pp; English.  
 XX  
 CC The sequence is of a Japanese cedar pollen allergen Cry j IIA  
 CC fragment. The protein and fragments can be used for diagnosis and  
 CC treatment of Japanese cedar pollinosis and to identify similar  
 CC sequences in other plants.  
 CC See also AAR53690-5.  
 XX  
 SQ Sequence 14 AA;

Query Match 28.8%; Score 30; DB 15; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
 I: |||:  
 Db 2 gyfsgnvi 9

RESULT 11  
 AAR97942  
 ID AAR97942 standard; peptide; 15 AA.  
 XX  
 AC AAR97942;  
 XX  
 DT 16-AUG-1996 (first entry)  
 XX  
 DE Japan cedar pollen mature allergen Cry j II amino acids 356-370.  
 XX  
 KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;  
 KW Sugi pollinosis; diagnosis; treatment.  
 XX  
 OS Cryptomeria japonica.  
 XX  
 PN JP08047392-A.  
 XX  
 PD 20-FEB-1996.  
 XX  
 PF 07-NOV-1994; 94JP-0297840.  
 XX  
 PR 26-MAY-1994; 94JP-0134868.  
 PR 05-NOV-1993; 93JP-0276773.  
 XX  
 PA (MEIP ) MEIJI MILK PROD CO LTD.  
 XX  
 DR WPI; 1996-166249/17.  
 XX  
 PT Japan cedar pollen allergen Cry j II epitope - comprises at least  
 PT part of specified 460 amino acid protein  
 XX  
 PS Claim 8; Fig 5; 17pp; Japanese.  
 XX  
 CC AAR97871-R97960 are overlapping peptides used for the epitope mapping  
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic  
 CC peptides of it are useful in the diagnosis, prevention and treatment  
 CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.  
 CC Significant regions of the allergen were identified using the  
 CC overlapping peptides of the full epitope derived from a Cry j II  
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200  
 CC (R978908) of the full mature 460 amino acid allergen are the most  
 CC allergenic of the 90 peptides tested.  
 XX  
 SQ Sequence 15 AA;

Query Match 28.8%; Score 30; DB 15; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
 I: |||:  
 Db 2 gyfsgnvi 9

RESULT 11  
 AAR97942  
 ID AAR97942 standard; peptide; 15 AA.  
 XX  
 AC AAR97942;  
 XX  
 DT 16-AUG-1996 (first entry)  
 XX  
 DE Japan cedar pollen mature allergen Cry j II amino acids 356-370.  
 XX  
 KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;  
 KW Sugi pollinosis; diagnosis; treatment.  
 XX  
 OS Cryptomeria japonica.  
 XX  
 PN JP08047392-A.  
 XX  
 PD 20-FEB-1996.  
 XX  
 PF 07-NOV-1994; 94JP-0297840.  
 XX  
 PR 26-MAY-1994; 94JP-0134868.  
 PR 05-NOV-1993; 93JP-0276773.  
 XX  
 PA (MEIP ) MEIJI MILK PROD CO LTD.  
 XX  
 DR WPI; 1996-166249/17.  
 XX  
 PT Japan cedar pollen allergen Cry j II epitope - comprises at least  
 PT part of specified 460 amino acid protein  
 XX  
 PS Claim 8; Fig 5; 17pp; Japanese.  
 XX  
 CC AAR97871-R97960 are overlapping peptides used for the epitope mapping  
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic  
 CC peptides of it are useful in the diagnosis, prevention and treatment  
 CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.  
 CC Significant regions of the allergen were identified using the  
 CC overlapping peptides of the full epitope derived from a Cry j II  
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200  
 CC (R978908) of the full mature 460 amino acid allergen are the most  
 CC allergenic of the 90 peptides tested.  
 XX  
 SQ Sequence 15 AA;

Query Match 28.8%; Score 30; DB 17; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
 I: |||:  
 Db 8 gyfsgnvi 15

RESULT 12  
 AAR97943  
 ID AAR97943 standard; peptide; 15 AA.  
 XX  
 AC AAR97943;  
 XX  
 DT 16-AUG-1996 (first entry)  
 XX  
 DE Japan cedar pollen mature allergen Cry j II amino acids 361-375.  
 XX  
 KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;  
 KW Sugi pollinosis; diagnosis; treatment.  
 XX  
 OS Cryptomeria japonica.  
 XX  
 PN JP08047392-A.  
 XX  
 PD 20-FEB-1996.  
 XX  
 PF 07-NOV-1994; 94JP-0297840.  
 XX  
 PR 26-MAY-1994; 94JP-0134868.  
 PR 05-NOV-1993; 93JP-0276773.  
 XX  
 PA (MEIP ) MEIJI MILK PROD CO LTD.  
 XX  
 DR WPI; 1996-166249/17.  
 XX  
 PT Japan cedar pollen allergen Cry j II epitope - comprises at least  
 PT part of specified 460 amino acid protein  
 XX  
 PS Claim 8; Fig 5; 17pp; Japanese.  
 XX  
 CC AAR97871-R97960 are overlapping peptides used for the epitope mapping  
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic  
 CC peptides of it are useful in the diagnosis, prevention and treatment  
 CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.  
 CC Significant regions of the allergen were identified using the  
 CC overlapping peptides of the full epitope derived from a Cry j II  
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200  
 CC (R978908) of the full mature 460 amino acid allergen are the most  
 CC allergenic of the 90 peptides tested.  
 XX  
 SQ Sequence 15 AA;

Query Match 28.8%; Score 30; DB 17; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
 I: |||:  
 Db 3 gyfsgnvi 10

RESULT 13  
 AAY57842  
 ID AAY57842 standard; peptide; 15 AA.  
 XX  
 AC AAY57842;  
 XX  
 DT 22-MAR-2000 (first entry)  
 XX  
 DE Human leptin receptor intracellular peptide EC-2 residues 473-487.

XX Human: leptin receptor; modulation; angiogenesis; wound healing;  
 KW regulation; ischaemia; leptin; tumour; metastasis; fat metabolism;  
 KW diabetic microangiopathy; neovascularisation.  
 XX  
 OS Homo sapiens.  
 PN WO959614-A1.  
 PD 25-NOV-1999.  
 XX  
 XX 20-MAY-1999; 99WO-US11209.  
 XX  
 PR 20-MAY-1998; 98US-0086354.  
 XX  
 PA (UYUA ) UNIV YALE.  
 XX  
 PI Sierra-Honigsmann RM;  
 XX  
 DR WPI; 2000-086576/07.  
 XX  
 PT Modulating an angiogenesis-inducing stimulus in wound healing, used in  
 PT treating, e.g. tumors -  
 XX  
 PS Example 1; Page 39; 89pp; English.  
 XX  
 CC A method has been developed for modulating a response in a subject to an  
 CC angiogenesis-inducing stimulus by administering an agent (I) that  
 CC modulates a leptin, or leptin receptor, mediated angiogenic response to  
 CC that stimulus. The method may be used to treat diseases controlled  
 CC by angiogenesis including wound healing, tumours and tumour metastasis,  
 CC diabetic microangiopathy, retinal neovascularisation, neovascularisation  
 CC of adipose tissue and fat metabolism, revascularisation of necrotic  
 CC tissue, enhancement or vascularisation in microvascular transplants and  
 CC ovarian follicle maturation. The present sequence represents a peptide  
 CC based on the human leptin receptor, given in an example from the present  
 CC invention.  
 XX  
 SQ Sequence 15 AA;

Query Match 28.8%; Score 30; DB 21; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 IPRHPNS 18  
 || || || ||  
 Db 4 ipsihpis 11

RESULT 14  
 AAM19282  
 ID AAM19282 standard; Protein; 18 AA.  
 AC AAM19282;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #5716 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00670.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 XX Claim 27; SEQ ID NO 24108; 487pp; English.  
 PS  
 XX The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
 XX  
 SQ Sequence 18 AA;  
 Query Match 28.8%; Score 30; DB 22; Length 18;  
 Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GFTEGHVVP 12  
 |::||:  
 Db 6 gwsghadlp 15  
 RESULT 15  
 AAM32006  
 ID AAM32006 standard; Protein; 18 AA.  
 AC AAM32006;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #6043 encoded by probe for measuring placental gene expression.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;



XX WPI; 2001-488897/53.  
DR  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX  
PS Claim 27; SEQ ID No 32275; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENP;  
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 18 AA;  
  
Query Match 28.8%; Score 30; DB 22; Length 18;  
Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 GFIEGHVWIP 12  
|::||:|  
Db 6 gwtvsgghadlp 15

Search completed: February 21, 2002, 17:16:11  
Job time: 191 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:14:40 ; Search time 12.47 Seconds  
(without alignments)  
34.287 Million cell updates/sec

Title: US-08-753-851-15

Perfect score: 104

Sequence: 1 RYGFEGHWIPRIHPNSI 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 106337

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	19	2	US-08-143-311B-13
2	32	30.8	18	2	US-08-644-271-16
3	31	29.8	19	1	US-08-492-599-3
4	31	29.8	19	5	PCT-US96-10455-3
5	30	28.8	10	3	US-08-159-339A-462
6	30	28.8	14	3	US-08-467-023-184
7	30	28.8	14	3	US-08-467-023-186
8	29	27.9	15	1	US-07-854-603-12
9	28.5	27.4	15	4	US-08-743-168B-20
10	28.5	27.4	15	5	PCT-US96-10435-20
11	28.5	27.4	19	2	US-08-374-652C-48
12	28	26.9	15	1	US-08-268-251-23
13	28	26.9	15	1	US-07-854-603-14
14	28	26.9	15	5	PCT-US93-01112-23
15	28	26.9	16	2	US-08-310-912A-54
16	28	26.9	16	3	US-08-841-089-54
17	28	26.9	16	4	US-09-301-085-54
18	28	26.9	16	5	PCT-US95-04570-54
19	28	26.9	16	5	PCT-US95-04589-54
20	28	26.9	19	2	US-08-436-420-16
21	27	26.0	15	1	US-07-854-603-10
22	27	26.0	17	1	US-08-787-547-27
23	27	26.0	19	1	US-08-732-751-23
24	27	26.0	19	2	US-08-807-332B-15
25	27	26.0	19	4	US-09-338-876-15
26	26.5	25.5	16	4	US-09-248-588-47
27	26	25.0	11	1	US-08-665-966-4

28 26 25.0 11 3 US-09-041-780-4 Sequence 4, Appli  
29 26 25.0 13 5 PCT-US95-09261-5 Sequence 5, Appli  
30 26 25.0 15 3 US-08-159-339A-1120 Sequence 1120, Ap  
31 26 25.0 16 2 US-07-888-943-3 Sequence 3, Appli  
32 26 25.0 16 2 US-08-100-691-2 Sequence 2, Appli  
33 26 25.0 18 1 US-07-872-644-15 Sequence 15, Appl  
34 26 25.0 18 1 US-08-297-494-15 Sequence 15, Appl  
35 26 25.0 18 1 US-08-297-510-15 Sequence 15, Appl  
36 26 25.0 18 1 US-08-479-532-15 Sequence 15, Appl  
37 26 25.0 18 1 US-08-455-526-15 Sequence 15, Appl  
38 26 25.0 18 1 US-08-455-526-15 Sequence 15, Appl  
39 26 25.0 18 2 US-08-522-269B-9 Sequence 9, Appli  
40 26 25.0 18 3 US-09-139-491-15 Sequence 15, Appl  
41 26 25.0 18 3 US-08-779-764A-34 Sequence 34, Appl  
42 26 25.0 18 4 US-09-294-923-9 Sequence 9, Appli  
43 26 25.0 18 5 PCT-US92-03222-15 Sequence 15, Appl  
44 26 25.0 19 1 US-08-492-599-2 Sequence 2, Appli  
45 26 25.0 19 2 US-08-031-538-44 Sequence 44, Appl

#### ALIGNMENTS

RESULT 1  
US-08-143-311B-13  
; Sequence 13, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311B-13

Query Match 100.0%; Score 104; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYGFIEGHVWIPRIHPNSI 19  
| | | | | | | | | | | | | | | | | | |  
Db 1 RYGFIEGHVWIPRIHPNSI 19

## RESULT 2

US-08-644-271-16  
; Sequence 16, Application US/08644271  
; Patent No. 5814478  
; GENERAL INFORMATION:  
; APPLICANT: Valenzuela, et al.  
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS  
; TITLE OF INVENTION: AND LIGANDS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill Road  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10591  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/644,271  
; FILING DATE: 10-MAY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 60/008,657  
; FILING DATE: 15-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cobert, Robert J.  
; REGISTRATION NUMBER: 36,108  
; REFERENCE/DOCKET NUMBER: REG 195A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 914-345-7400  
; TELEFAX: 914-345-7721  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Other  
; LOCATION: 1..18  
; OTHER INFORMATION: No. 5814478enclosure for this antibody  
; OTHER INFORMATION: Is 52307K.  
US-08-644-271-16

Query Match 30.8%; Score 32; DB 2; Length 18;  
Best Local Similarity 30.8%; Pred. No. 29;  
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 IEQHVWIPRIHPN 17  
: : : : : | | | | |  
Db 2 LPSELLDLRLHPN 14

## RESULT 3

US-08-492-599-3

; Sequence 3, Application US/08492599  
; Patent No. 5726050  
; GENERAL INFORMATION:  
; APPLICANT: Rich, Alexander  
; APPLICANT: Herbert, Alan  
; TITLE OF INVENTION: Z-DNA BINDING PROTEIN AND APPLICATIONS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kevin M. Farrell, P.C.  
; STREET: P.O. Box 999  
; CITY: York Harbor  
; STATE: ME  
; COUNTRY: US  
; ZIP: 03911  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/492,599  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farrell, Kevin M.  
; REGISTRATION NUMBER: 35,505  
; REFERENCE/DOCKET NUMBER: MIT-7011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (207) 363-0538  
; TELEFAX: (207) 363-0528  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-492-599-3  
  
Query Match 29.8%; Score 31; DB 1; Length 19;  
Best Local Similarity 42.9%; Pred. No. 45;  
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
QY 1 RYGFIEGHVWIPRI 14  
| | | | | : | :  
Db 2 RYPTIVNHPKVGRI 15  
  
RESULT 4  
PCT-US96-10455-3  
; Sequence 3, Application PC/TUS9610455  
; GENERAL INFORMATION:  
; APPLICANT: Massachusetts Institute of Technology  
; TITLE OF INVENTION: Z-DNA BINDING PROTEIN AND APPLICATIONS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kevin M. Farrell, P.C.  
; STREET: P.O. Box 999  
; CITY: York Harbor  
; STATE: ME  
; COUNTRY: US  
; ZIP: 03911  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10455  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/492,599  
FILING DATE: 20-JUNE-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farrell, Kevin M.  
REGISTRATION NUMBER: 35,505  
REFERENCE/DOCKET NUMBER: MIT-7011 WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (207) 363-0558  
TELEFAX: (207) 363-0528  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US96-10455-3

Query Match 29.8%; Score 31; DB 5; Length 19;  
Best Local Similarity 42.9%; Pred. No. 45;  
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 RYGFIEGHVVIPRI 14  
DB 2 RYPFIVNHPKVGRI 15

RESULT 5  
US-08-159-339A-462

Sequence 462, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and Their  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/936,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-0050300S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 462:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-462

Query Match 28.8%; Score 30; DB 3; Length 10;  
Best Local Similarity 44.4%; Pred. No. 32;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 IECHVWIPR 13  
DB 2 VNGHQVLPK 10

RESULT 6

US-08-467-023-184  
Sequence 184, Application US/08467023  
Patent No. 6090386  
GENERAL INFORMATION:  
APPLICANT: Griffith, Irwin J.;  
APPLICANT: Pollock, Joanne;  
APPLICANT: Bond, Julian F.;  
APPLICANT: Garman, Richard D;  
APPLICANT: Kuo, Mei-Chang;  
APPLICANT: Yeung, Siu-mei H.;  
APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powers, Steven P.  
TITLE OF INVENTION: Allergenic proteins And Peptides From  
TITLE OF INVENTION: Japanese Cedar Pollen  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,023  
FILING DATE: June 6, 1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane E. Remillard  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 184:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-467-023-184

Query Match 28.8%; Score 30; DB 3; Length 14;  
Best Local Similarity 50.0%; Pred. No. 47;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
I: III;  
Db 2 GYFSGHVI 9

## RESULT 7

US-08-467-023-186  
; Sequence 186, Application US/08467023  
; Patent No. 6090386  
; GENERAL INFORMATION:  
; APPLICANT: Griffith, Irwin J.;  
; APPLICANT: Pollock, Joanne;  
; APPLICANT: Bond, Julian F.;  
; APPLICANT: Garman, Richard D;  
; APPLICANT: Kuo, Mei-Chang;  
; APPLICANT: Yeung, Siu-mei H.;  
; APPLICANT: Brauer, Andrew;  
; APPLICANT: Exley, Mark A.;  
; APPLICANT: Powers, Steven P.  
; TITLE OF INVENTION: Allergenic Proteins And Peptides From  
; TITLE OF INVENTION: Japanese Cedar Pollen  
; NUMBER OF SEQUENCES: 261  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
; STREET: 610 Lincoln St  
; CITY: Waltham  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,023  
; FILING DATE: June 6, 1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/350,225  
; FILING DATE: December 6, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane E. Remillard  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 186:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-467-023-186

Query Match 28.8%; Score 30; DB 3; Length 14;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
I: III;  
Db 2 GYFSGHVI 9

## RESULT 8

US-07-854-603-12  
; Sequence 12, Application US/07854603  
; Patent No. 5637492

; GENERAL INFORMATION:  
; APPLICANT: Dawson, Keith M  
; APPLICANT: Edwards, Richard M  
; APPLICANT: Forman, Joan M  
; TITLE OF INVENTION: Activatable fibrinolytic and  
; TITLE OF INVENTION: anti-thrombotic proteins  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dr. John J. McDonnell  
; STREET: Ten South Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/854,603  
; FILING DATE: 19901207  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McDonnell, John J  
; REGISTRATION NUMBER: 26,949  
; REFERENCE/DOCKET NUMBER: 92,338  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; TELEX: 910-221-5317  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-854-603-12

Query Match 27.9%; Score 29; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. No. 75;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GFIEGHVV 10  
I: I: I I  
Db 5 GYIDGRVV 12

## RESULT 9

US-08-743-168B-20  
; Sequence 20, Application US/08743168B  
; Patent No. 6271015  
; GENERAL INFORMATION:  
; APPLICANT: Gilula, No. 6271015ton B  
; APPLICANT: Cravatt, Benjamin F  
; APPLICANT: Lerner, Richard A  
; TITLE OF INVENTION: FATTY-ACID AMIDE HYDROLASE  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute  
; STREET: 10550 No. 6271015th Torrey Pines Road  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/743,168B

```

; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US 08/489,535
; FILING DATE: 12-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 485.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-743-168B-20

Query Match 27.4%; Score 28.5; DB 4; Length 15;
Best Local Similarity 50.0%; Pred. No. 91;
Matches 6; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 7 GHVVIPRIHPNS 18
Db 5 GHTLIPL-PNN 15

RESULT 10
PCT-US96-10435-20
; Sequence 20, Application PC/TUS9610435
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: CIS-9, 10-OCTADECENOAMIDASE
; NUMBER OF SEQUENCES: 32
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10435
; FILING DATE: 12-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/489,535
; FILING DATE: 12-JUN-1995
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; PCT-US96-10435-20

Query Match 27.4%; Score 28.5; DB 5; Length 15;
Best Local Similarity 50.0%; Pred. No. 91;
Matches 6; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 7 GHVVIPRIHPNS 18
Db 5 GHTLIPL-PNN 15

RESULT 11
US-08-374-652C-48
; Sequence 48, Application US/08374652C
; Patent No. 5834286

; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTATE DEGRADING
; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058
; FILING DATE: 27-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925,401
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1050.071001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-374-652C-48

Query Match 27.4%; Score 28.5; DB 2; Length 19;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 RYGF-EGHVVIP 12
Db 2 RYHLWNGETVVP 14

RESULT 12
US-08-268-251-23
; Sequence 23, Application US/08268251
; Patent No. 5585475
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
```

```
;
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/268,251
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-268-251-23

Query Match 26.9%; Score 28; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 FIEGHV 9
DB 10 YLEGHV 15

RESULT 13
US-07-854-603-14
; Sequence 14, Application US/07854603
; Patent No. 5637492
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Edwards, Richard M
; APPLICANT: Forman, Joan M
; TITLE OF INVENTION: Activatable fibrinolytic and
; TITLE OF INVENTION: anti-thrombotic proteins
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,603
; FILING DATE: 19901207
; CLASSIFICATION:
```

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELE: 910-221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-854-603-14

Query Match 26.9%; Score 28; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GFIEGHV 10
DB 5 GYIDGRIV 12

RESULT 14
PCT-US93-01112-23
; Sequence 23, Application PC/TUS9301112
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01112
; FILING DATE: 19930208
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-01112-23

Query Match 26.9%; Score 28; DB 5; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
```

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 FIEGHV 9  
:||||  
Db 10 YLEGHV 15

## RESULT 15

US-08-310-912A-54  
; Sequence 54, Application US/08310912A  
; Patent No. 5981730  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindrinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION  
; NUMBER OF SEQUENCES: 208  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/310,912A  
; FILING DATE: September 22, 1994  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/227,360  
; FILING DATE: April 13, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lech, Karen F.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/254001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 100254  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-310-912A-54

Query Match 26.9%; Score 28; DB 2; Length 16;  
Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 13 RIHPNSI 19  
:||||  
Db 4 RVHPSKI 10

Search completed: February 21, 2002, 17:16:30  
Job time: 110 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000, Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:18:50 ; Search time 12.71 Seconds  
(without alignments)  
65.926 Million cell updates/sec

Title: US-08-753-851-16

Perfect score: 61

Sequence: 1 LTYNTSQYDTY 11

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 1324

Minimum DB seq length: 0

Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	32.8	9	2 S70334	endosperm protein,
2	20	32.8	11	2 A60656	perisulfakinin - A
3	19	31.1	9	2 PW0002	chlorophyll a/b-bi
4	18	29.5	10	2 PT0243	Ig heavy chain CRD
5	18	29.5	11	2 A32428	amine oxidase (cop
6	17	27.9	7	2 PC1316	large granule L3 c
7	17	27.9	8	2 S43971	tumor-associated a
8	17	27.9	8	2 S43972	tumor-associated a
9	17	27.9	8	2 S11078	glucose-6-phosphat
10	17	27.9	9	2 A24244	adipokinetic hormo
11	17	27.9	10	2 PG0037	alpha-peptide/alg
12	17	27.9	10	2 A47364	placental lactogen
13	17	27.9	11	2 S23373	T-cell receptor al
14	16	26.2	9	2 A28924	fructose-bisphosph
15	16	26.2	10	2 S39374	mannose receptor -
16	16	26.2	10	2 PT0291	Ig heavy chain CRD
17	16	26.2	11	1 GMROL	leucosulfakinin -
18	16	26.2	11	2 PH0919	T-cell receptor be
19	15	24.6	8	2 PT0311	Ig heavy chain CRD
20	15	24.6	9	2 S36850	Ig heavy chain v r
21	15	24.6	10	1 GMROL2	leucosulfakinin-II
22	15	24.6	10	2 S74176	gluconokinase (EC
23	15	24.6	10	2 S62208	polyferredoxin - M
24	15	24.6	10	2 C61440	polygalacturonase
25	15	24.6	10	2 B60656	leucosulfakinin II
26	15	24.6	10	2 PT0215	T-cell receptor be
27	15	24.6	11	2 S05002	corazonin - Americ
28	15	24.6	11	2 A40795	glycoprotein H-a -
29	15	24.6	11	2 PT0214	T-cell receptor be

68kDa neurofilamen  
beta-D-galactosida  
phenol 2-monooxyge  
hypothalamic hepta  
disstimilatory sulf  
calsequestrin, car  
hypothetical prote  
callipmrfamide 1 -  
Ig heavy chain CRD  
enamelin 1 - bovin  
pyrimidine synthes  
T-cell receptor ga  
peptidylglycine mo  
T-cell receptor be  
hypertrahalosemic/  
probable glucose-6

#### ALIGNMENTS

RESULT 1

S70334

endosperm protein, 45K - rye (fragment)

C:Species: Secale cereale (rye)

C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 17-Apr-1998

C:Accession: S70334

R;Rocher, A.; Calero, M.; Soriano, F.; Mendez, E.

Biochim. Biophys. Acta 1295, 13-22, 1996

A:Title: Identification of major rye secalins as coeliac immunoreactive proteins.

A:Reference number: S70327; MUID:96283789

A:Accession: S70334

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <ROC>

Query Match 32.8%; Score 20; DB 2; Length 9;  
Best Local Similarity 57.1%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 TSQYDTY 11

|||

Db 1 TSIYQXY 7

RESULT 2

A60656

perisulfakinin - American cockroach

C:Species: Periplaneta americana (American cockroach)

C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 11-Jul-1997

C:Accession: A60656

R;Veenstra, J.A.

Neuropeptides 14, 145-149, 1989

A:Title: Isolation and structure of two gastrin/CCK-like neuropeptides from the Ameri

A:Reference number: A60656; MUID:90137190

A:Accession: A60656

A:Molecule type: protein

A:Residues: 1-11 <VEE>

C:Comment: This neuropeptide stimulates hindgut contractions.

C:Keywords: amidated carboxyl end; neuropeptide; sulfoprotein

F:6/Binding site: sulfate (Tyr) (covalent) #status experimental

F:11/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match

Best Local Similarity 60.0%; Pred. No. 1e+03;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 QYDTY 11

|||

Db 2 QFDDY 6



C:Superfamily: unassigned animal peptides

Query Match 27.9%; Score 17; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 NTSQ 7  
II: I  
Db 4 NTAQ 7

RESULT 9  
S11078  
glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) - yeast (*Pichia jadinii*) (fragment)  
C:Species: *Pichia jadinii*, *Candida utilis*  
C:Date: 30-Jun-1991 #sequence\_revision 30-Sep-1991 #text\_change 05-Aug-1994  
C:Accession: S11078  
R:Egestad, B.; Estonius, M.; Danielsson, O.; Persson, B.; Cederlund, E.; Kaiser, R.; Hol  
FEBS Lett. 269, 194-196, 1990  
A:Title: Fast atom bombardment mass spectrometry and chemical analysis in determinations  
A:Reference number: S11074; MUID:90353571  
A:Accession: S11078  
A:Molecule type: protein  
A:Residues: 1-8 <ECE>  
A:Note: the source is designated as *Pichia jadinii*  
C:Keywords: acetylated amino end; oxidoreductase; pentose phosphate pathway  
F:1/Modified site: acetylated amino end (Ser) #status experimental

Query Match 27.9%; Score 17; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 YDTY 11  
II: I  
Db 2 YDSF 5

RESULT 10  
A24244  
adipoketic hormone - bollworm  
N:Alternate names: Hez-AKH  
C:Species: *Heliothis zea* (bollworm, corn earworm, tomato fruitworm)  
C:Date: 31-Mar-1988 #sequence\_revision 23-Mar-1995 #text\_change 31-Oct-1997  
C:Accession: A24244  
R:Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Holman, G.M.; Wagner, R.M.; Ridgway  
Biochem. Biophys. Res. Commun. 135, 622-628, 1986  
A:Title: Isolation and primary structure of a peptide from the corpora cardiaca of *Heliothis zea*  
A:Reference number: A24244; MUID:86186794  
A:Accession: A24244  
A:Molecule type: protein  
A:Residues: 1-9 <JAF>  
C:Superfamily: adipoketic hormone  
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyroglutamic carboxylic acid (Gln) #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 27.9%; Score 17; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTYNTS 6  
II: I  
Db 2 LFTSS 7

RESULT 11  
PU0037  
alpha-peptide/alginate lyase fused protein 1 - *Pseudomonas* sp. (fragment)  
C:Species: *Pseudomonas* sp.  
C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 12-Sep-1997

C:Accession: PU0037  
R:Fujiyama, K.; Maki, H.; Kinoshita, S.; Yoshida, T.  
submitted to JIPID, October 1994  
A:Description: High gene expression in *E. coli* of recombinant alginate lyase as a fus  
A:Reference number: PU0037  
A:Accession: PU0037  
A:Molecule type: protein  
A:Residues: 1-10 <FUJ>  
C:Genetics: -  
A:Gene: lacZ'/aly

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 3.2e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 NTSQY 8  
II: I  
Db 4 NSSXY 8

RESULT 12  
A47364  
placental lactogen-I precursor - mouse (fragment)  
C:Species: *Mus sp.* (mouse)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 12-Aug-1996  
C:Accession: A47364  
R:Shida, M.M.; Ng, Y.K.; Soares, M.J.; Linzer, D.I.  
Mol. Endocrinol. 7, 181-188, 1993  
A:Title: Trophoblast-specific transcription from the mouse placental lactogen-I gene  
A:Reference number: A47364; MUID:93225959  
A:Accession: A47364  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-10 <RES>  
A:Cross-references: GB:S58124; NID:g299449

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTYNTS 6  
II: I  
Db 3 LTLNLS 8

RESULT 13  
S23373  
T-cell receptor alpha chain J region - human (fragment)  
C:Species: *Homo sapiens* (man)  
C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
C:Accession: S23373  
R:Pluschke, G.; Ricken, G.; Taube, H.; Kroninger, S.; Melchers, I.; Peter, H.H.; Eich  
Eur. J. Immunol. 21, 2749-2754, 1991  
A:Title: Biased T cell receptor V(alpha) region repertoire in the synovial fluid of r  
A:Reference number: S23364; MUID:92037820  
A:Accession: S23373  
A:Status: preliminary; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-11 <PLU>  
A:Cross-references: EMBL:X58168  
C:Keywords: T-cell receptor

Query Match 27.9%; Score 17; DB 2; Length 11;  
Best Local Similarity 75.0%; Pred. No. 3.5e+03;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 SQYD 9  
II: I  
Db 4 SNYD 7

## RESULT 14

A28924

fructose-bisphosphate aldolase (EC 4.1.2.13) B, hepatic - rabbit (fragment)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 07-Feb-1997  
 C:Accession: A28924  
 R:Lacko, A.G.; Brox, L.W.; Gracy, R.W.; Horecker, B.L.  
 J. Biol. Chem. 245, 2140-2141, 1970  
 A:Title: The carboxyl-terminal structure of rabbit liver aldolase (aldolase B).  
 A:Reference number: A28924; MUID:70166720  
 A:Accession: A28924  
 A:Molecule type: protein  
 A:Residues: 1-9 <LAK>  
 C:Keywords: aldehyde-lyase; carbon-carbon lyase; gluconeogenesis; glycolysis; liver; per

Query Match 26.2%; Score 16; DB 2; Length 9;

Best Local Similarity 75.0%; Pred. No. 2.2e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TSOY 8

| | |

Db 6 TSSY 9

## RESULT 15

S39374

mannose receptor - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 18-Feb-1994 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
 C:Accession: S39374  
 R:Messner, D.J.  
 Arch. Biochem. Biophys. 306, 391-401, 1993  
 A:Title: The mannose receptor and the cation-dependent form of mannose 6-phosphate recep  
 A:Reference number: S39374; MUID:94029009  
 A:Accession: S39374  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <MES>  
 C:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re  
 C:Keywords: membrane protein; receptor

Query Match 26.2%; Score 16; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 4.8e+03;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTYN 4

| | |

Db 5 LIYN 8

Search completed: February 21, 2002, 17:20:40

Job time: 110 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:20:26 ; Search time 10.02 Seconds  
(without alignments)  
40.251 Million cell updates/sec

Title: US-08-753-851-16  
Perfect score: 61  
Sequence: 1 LTYNTSQYDTY 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 388

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	23	37.7	10	UH05_RAT	P56573 rattus norv
2	20	32.8	11	LSKP_PERAM	P36885 periplaneta
3	17	27.9	9	AL11_CARMA	P81814 carcinus ma
4	17	27.9	10	UPA4_HUMAN	P30090 homo sapien
5	16	26.2	9	FARP_CALSI	P38495 callinectes
6	16	26.2	11	LSK1_LEUMA	P04428 leucophaea
7	15	24.6	9	NSK1_SARBU	P41492 sarcophaga
8	15	24.6	10	GS09_BACSU	P80243 bacillus su
9	15	24.6	10	LSK2_LEUMA	P09039 leucophaea
10	15	24.6	11	CORZ_PERAM	P11496 periplaneta
11	14	23.0	6	UN06_CLOPA	P81351 clostridium
12	14	23.0	7	HV7_PIG	P01153 sus scrofa
13	14	23.0	7	EAR1_CALVO	P41856 calliphora
14	14	23.0	9	UH21_HUMAN	P40929 homo sapien
15	14	23.0	10	HTF_HELZE	P16353 heliothis z
16	14	23.0	10	UH31_HUMAN	P40930 homo sapien
17	14	23.0	11	RANP_RANPI	P08951 rana pipien
18	13	21.3	6	ASP2_LACSN	P82655 lactobacill
19	13	21.3	7	ALL7_CYDPO	P82158 cydia pomon
20	13	21.3	7	MNP1_LEPDE	P42984 leptinotars
21	13	21.3	8	AL11_CARMA	P42984 carcinus ma
22	13	21.3	8	ALL1_CYDPO	P82152 cydia pomon
23	13	21.3	8	ALL5_CALVO	P41841 calliphora
24	13	21.3	8	ALL5_CYDPO	P82156 cydia pomon
25	13	21.3	8	ALL6_CYDPO	P82157 cydia pomon
26	13	21.3	9	ULAD_HUMAN	P31929 homo sapien
27	13	21.3	10	COXO_RAT	P80432 rattus norv
28	13	21.3	10	ESTA_SCHGA	P81012 schizaphis
29	13	21.3	10	NO40_TOBAC	P55962 nicotiana t
30	13	21.3	10	SP34_DICMU	P81545 dictyosteli
31	13	21.3	10	TKN1_SCYCA	P08608 scyllorhinu
32	13	21.3	10	THOF_AEDAE	P19425 aedes aegypt
33	13	21.3	11	PKC1_CARMO	P82684 carausius m

34	13	21.3	11	1	T2P1_PROVU	P31031 proteus vul
35	12	19.7	7	1	ALL4_CARMA	P81807 carcinus ma
36	12	19.7	8	1	AKH_MELML	P25423 melolontha
37	12	19.7	8	1	AKH_TARAT	P14595 tabanus atr
38	12	19.7	8	1	AL12_CARMA	P81815 carcinus ma
39	12	19.7	8	1	AL18_CARMA	P81821 carcinus ma
40	12	19.7	8	1	HTF2_PERAM	P04549 periplaneta
41	12	19.7	8	1	ORMY_ORCLI	P82455 orconectes
42	12	19.7	8	1	UPAA_HUMAN	P30096 homo sapien
43	12	19.7	9	1	HUTU_KLEAE	P12381 klebsiella
44	12	19.7	9	1	PGLR_DIAAB	P81179 diaprepes a
45	12	19.7	10	1	AL19_CARMA	P81822 carcinus ma

ALIGNMENTS

RESULT 1

UH05\_RAT STANDARD; PRT; 10 AA.  
AC P56573;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF HEART TISSUE (SPOT P5) (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=WISTAR; TISSUE=Heart;  
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,  
RA Jungblut P.R.;  
RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 8.3, ITS MW IS: 30 KDA.  
FT UNSURE 9 9 OR P.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1269 MW; 00CBA4BB46C5BABB CRC64;

Query Match 37.7%; Score 23; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 86;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 SQYDTY 11  
DB 5 SQYDGF 10

RESULT 2

LSKP\_PERAM STANDARD; PRT; 11 AA.  
AC P36885;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE PERISULFAKININ (PEA-SK-I).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blattoidea; Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Corpora cardiaca;  
RC MEDLINE=90137190; PubMed=2615921;  
RA Veenstra J.A.;  
RT "Isolation and structure of two gastrin/CCK-like neuropeptides from  
RT the American cockroach homologous to the leucosulfrakinins.";  
RL Neuropeptides 14:145-149(1989).  
CC -!- FUNCTION: STIMULATES HINDGUT CONTRACTIONS.

CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 DR PIR; A60656; A60656.  
 DR InterPro: IPR001651; Gastrin.  
 DR PROSITE: PS00259; GASTRIN; 1.  
 KW Hormone; Amidation; Sulfation.  
 FT MOD\_RES 6 6 SULFATION.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1445 MW; 8B4E0680EB6B5AAA CRC64;

Query Match 32.8%; Score 20; DB 1; Length 11;  
 Best Local Similarity 60.0%; Pred. No. 3.3e+02;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 QYDTY 11  
 I I I  
 Db 2 QPDDY 6

RESULT 3  
 ID AL11\_CARMA STANDARD; PRT; 9 AA.  
 AC P81814;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE CARCINUSTATIN 11.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 Eur. J. Biochem. 250:727-734 (1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 927 MW; 832D79DCB46D861 CRC64;

Query Match 27.9%; Score 17; DB 1; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 TSQY 8  
 I I I  
 Db 2 TGQY 5

RESULT 4  
 ID UPA4\_HUMAN STANDARD; PRT; 10 AA.  
 AC P30090;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 12) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=Q3092937; PubMed=1459097;

RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing.";  
 RL Electrophoresis 13:707-714 (1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.8, ITS MW IS: 40.5 KDA.  
 CC -!- MISCELLANEOUS: THIS SPOT IS ON A POSITION THOUGHT TO BE THAT OF  
 CC ZN-ALPHA-2 GLYCOPROTEIN, BUT IT DOES NOT CORRESPOND TO THAT  
 CC PROTEIN.  
 DR SWISS-2DPAGE; P30090; HUMAN.  
 FT NON\_TER 1 1  
 FT UNSURE 4 4  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1067 MW; 269492EB05A1A457 CRC64;

Query Match 27.9%; Score 17; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 NTSQ 7  
 I I I  
 Db 4 NTSE 7

RESULT 5  
 ID FARP\_CALSI STANDARD; PRT; 9 AA.  
 AC P38495;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE.  
 OS Callinectes sapidus (Blue crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Callinectes.  
 OX NCBI\_TaxID=6763;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92270479; PubMed=1815216;  
 RA Krajncak K.G.;  
 RT "The identification and structure-activity relations of a  
 RT cardioactive FMRFamide-related peptide from the blue crab Callinectes  
 RT sapidus.";  
 RL Peptides 12:1295-1302 (1991).  
 CC -!- FUNCTION: CARDIOACTIVE PEPTIDE.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 1159 MW; 134F0729D5A4045B CRC64;

Query Match 26.2%; Score 16; DB 1; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 YNTS 6  
 I I I  
 Db 2 YNRS 5

RESULT 6  
 ID LSK1\_LEUMA STANDARD; PRT; 11 AA.  
 AC P04428;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE LEUCOSULFAKININ-I (LSK-I).  
 OS Leucophaea maderae (Madeira cockroach).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=8631588; PubMed=3749893;  
 RA Nachman R.J., Holman G.M., Haddon W.F., Ling N.;  
 RT "Leucosulfakinin, a sulfated insect neuropeptide with homology to  
 gastrin and cholecystokinin.";  
 RL Science 234:71-73(1986).  
 CC -!- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF  
 THE COCKROACH HINGUT. INHIBITS MUSCLE CONTRACTION OF HINDGUT.  
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 CC PIR: A01622; GMR0L.  
 DR InterPro: IPR001651; Gastrin.  
 DR PROSITE: PS00259; GASTRIN; 1.  
 KW Hormone; Amidation; Sulfation.  
 FT MOD\_RES 6 6 SULFATION.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1459 MW; 7E4E0680E8B5AAB CRC64;  
 Query Match 26.2%; Score 16; DB 1; Length 11;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+03;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 7 QYDTY 11  
 Db 2 QFEDY 6  
 RESULT 7  
 ID NSK1\_SARBU STANDARD; PRT; 9 AA.  
 AC P41492;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE NEOSULFALKININ-I (NEB-SK-I).  
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Sarcophagidae; Sarcophaga.  
 OX NCBI\_TaxID=7385;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Head;  
 RX MEDLINE=93083101; PubMed=1360367;  
 RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;  
 RT "Isolation and primary structure of two sulfakinin-like peptides from  
 the fleshfly, Neobellieria bullata.";  
 RL Comp. Biochem. Physiol. 103C:135-142(1992).  
 CC -!- FUNCTION: MYOTROPIC PEPTIDE.  
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 DR InterPro: IPR001651; Gastrin.  
 DR PROSITE: PS00259; GASTRIN; 1.  
 KW Neuropeptide; Amidation; Sulfation.  
 FT MOD\_RES 4 4 SULFATION (POTENTIAL).  
 FT MOD\_RES 9 9 AMIDATION (POTENTIAL).  
 SQ SEQUENCE 9 AA; 1187 MW; 8B0A0691E86B5AAA CRC64;

Query Match 24.6%; Score 15; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 8 YDTY 11  
 Db 1 FDY 4  
 RESULT 8

GS09\_BACSU STANDARD; PRT; 10 AA.  
 ID GS09\_BACSU  
 AC P80243;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE GENERAL STRESS PROTEIN 9 (GSP9) (FRAGMENT).  
 GN YZF.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=IS58;  
 RX MEDLINE=94282319; PubMed=8012595;  
 RA Voelker U., Engelmann S., Maul B., Riethdorf S., Voelker A.,  
 RA Schmid R., Mach H., Hecker M.;  
 RT "Analysis of the induction of general stress proteins of Bacillus  
 subtilis.";  
 RL Microbiology 140:741-752(1994).  
 CC -!- INDUCTION: BY HEAT SHOCK, SALT STRESS, OXIDATIVE STRESS, GLUCOSE  
 LIMITATION AND OXYGEN LIMITATION.  
 CC Subtilisin; BGL9022; yzf.  
 KW Heat shock; Complete proteome.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1168 MW; 99766442DSA2C05A CRC64;

Query Match 24.6%; Score 15; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 2.4e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 SQYD 9  
 Db 6 SVYD 9

RESULT 9  
 ID LSK2\_LEUMA STANDARD; PRT; 10 AA.  
 AC P09039;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE LEUCOSULFALKININ-II (LSK-II).  
 OS Leucophaea maderae (Madeira cockroach), and  
 OS Periplaneta americana (American cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988, 6978;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=L.maderae;  
 RX MEDLINE=87048769; PubMed=3778455;  
 RA Nachman R.J., Holman G.M., Cook B.J., Haddon W.F., Ling N.;  
 RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with  
 homology to cholecystokinin and gastrin.";  
 RL Biochem. Biophys. Res. Commun. 140:357-364(1986).  
 RN [2]  
 RP SEQUENCE.

RC SPECIES=P.americana; TISSUE=Corpora cardiaca;  
 RX MEDLINE=90137190; PubMed=2615921;  
 RA Veenstra J.A.;  
 RT "Isolation and structure of two gastrin/CCK-like neuropeptides from  
 the American cockroach homologous to the leucosulfakinins.";  
 RL Neuropeptides 14:145-149(1989).  
 CC -!- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF  
 THE COCKROACH HINGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.  
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 DR PIR: A26335; GMR0L2.  
 DR PIR: B60656; B60656.

```

DR InterPro: IPR001651; Gastrin.
KW PROSITE: PS00259; GASTRIN; 1.
KW Hormone; Amidation; Sulfation.
FT MOD_RES 1 5 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 5 SULFATION (IN L-MADERAE, BUT NOT IN
FT MOD_RES 10 10 P-AMERICANA).
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1255 MW; 984F5391E8685AAA CRC64;

Query Match 24.6%; Score 15; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 2.4e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 QYDTY 11
DB 1 QSDDY 5

RESULT 10
CORZ_PERAM STANDARD; PRT; 11 AA.
ID P11496;
AC P11496;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE CORAZONIN.
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=89325572; PubMed=2753132;
RA Veenstra J.A.;
RT "Isolation and structure of corazonin, a cardioactive peptide from
RT the American cockroach.";
RL FEBS Lett. 250:231-234(1989).
CC -1- FUNCTION: CARDIOACTIVE PEPTIDE. CORAZONIN IS PROBABLY INVOLVED
CC IN THE PHYSIOLOGICAL REGULATION OF THE HEART BEAT.
DR PIR: S05002; S05002.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 11 11 AMIDATION.
SQ SEQUENCE 11 AA; 1387 MW; C7CFF32D6415AB46 CRC64;

Query Match 24.6%; Score 15; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 2.7e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TSOY 8
DB 2 TFOY 5

RESULT 11
UN06_CLOPA STANDARD; PRT; 6 AA.
ID UN06_CLOPA
AC P81351;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE UNKNOWN PROTEIN CP 6 FROM 2D-PAGE (FRAGMENT).
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=46;

```

```

RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC NON_TER 6
FT NON_TER 6
SQ SEQUENCE 6 AA; 657 MW; 605B1DC1A45A8000 CRC64;

Query Match 23.0%; Score 14; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 NTSQ 7
DB 2 NTAE 5

RESULT 12
HY7_PIG STANDARD; PRT; 7 AA.
ID HY7_PIG
AC P01153;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE HYPOTHALAMIC HEPTAPEPTIDE.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=81213980; PubMed=6263778;
RA Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
RA Safran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
RT "Isolation, structure and synthesis of a heptapeptide with in vitro
RT ACTH-releasing activity from porcine hypothalamus.";
RL Horm. Metab. Res. 13:228-232(1981).
DR PIR: A01417; NYPG7.
SQ SEQUENCE 7 AA; 957 MW; 632B45B1FB5059A0 CRC64;

Query Match 23.0%; Score 14; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 YDTY 11
DB 3 YHSY 6

RESULT 13
FARL_CALVO STANDARD; PRT; 9 AA.
ID FARL_CALVO
AC P41856;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE CALLIPWRFAMIDE 1.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of Phe-Met-Arg-Phe-NH2

```



RT neuropeptides (designated callifmrfamides) from the blowfly  
 RT Calliphora vomitoria.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
 CC -!- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED  
 CC SALIVARY GLAND OF CALLIPHORA.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR: A41978; A41978.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 1169 MW; 29D00699CAB6C6C7 CRC64;

Query Match 23.0%; Score 14; DB 1; Length 9;

Best Local Similarity 60.0%; Pred. No. 1e+05; 2; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 2;

QY 5 TSQYD 9

|||

Db 1 TPQD 5

#### RESULT 14

UHA2\_HUMAN STANDARD; PRT; 9 AA.  
 AC P40929;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF HEART (SPOT 5603) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.0, ITS MW IS: 55.3 KDA.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1104 MW; 8874B1BB5B01B2CA CRC64;

Query Match 23.0%; Score 14; DB 1; Length 9;

Best Local Similarity 16.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;  
 Matches 1; Conservative 4; Mismatches 1;

QY 3 YNTSQY 8

:::|

Db 1 FQVTEY 6

#### RESULT 15

HTF\_HELZE STANDARD; PRT; 10 AA.  
 AC P16353;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE HYPERTREHALOSAEMIC HORMONE (HEZ-HRTH).  
 OS Heliothis zea (Corn earworm) (Bollworm).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Noctuoidea; Noctuidae; Heliothinae; Helicoverpa.  
 OX NCBI\_TaxID=7113;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Corpora cardiaca;

RX MEDLINE=88326324; PubMed=3415690;  
 RA Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Bird T.G.,  
 RA Tseng C.M., Zhang Y.S., Hayes D.K.;  
 RT "Isolation and primary structure of a neuropeptide hormone from  
 RT Heliothis zea with hypertrehalosemic and adipokinetic activities.";  
 RL Biochem. Biophys. Res. Commun. 155:344-350(1988).  
 CC -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT  
 CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS  
 CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).  
 CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.  
 DR PIR: A31571; A31571.  
 DR InterPro: IPR002047; AKH.  
 DR PROSITE: PS00256; AKH; 1.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 1 1  
 FT MOD\_RES 10 10  
 SQ SEQUENCE 10 AA; 1096 MW; 8E70367865A5B9D1 CRC64;

Query Match 23.0%; Score 14; DB 1; Length 10;

Best Local Similarity 40.0%; Pred. No. 3.6e+03; 0; Gaps 0;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0;

QY 1 LTYNT 5

||||

Db 2 LTFSS 6

Search completed: February 21, 2002, 17:23:23  
 Job time: 177 sec

15

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 21, 2002, 17:20:06 ; Search time 22.04 Seconds  
(without alignments)  
73.003 Million cell updates/sec

Title: US-08-753-851-16

Perfect score: 61

Sequence: 1 LTYNTSQDYRY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 1277

Minimum DB seq length: 0

Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL17.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	22	36.1	11	7 077899	077899 oreochromis
2	22	36.1	11	7 077900	077900 oreochromis
3	22	36.1	11	7 077901	077901 oreochromis
4	22	36.1	11	7 077902	077902 oreochromis
5	22	36.1	11	7 077903	077903 oreochromis
6	22	36.1	11	7 077904	077904 oreochromis
7	22	36.1	11	7 077905	077905 oreochromis
8	22	36.1	11	7 077916	077916 oreochromis
9	22	36.1	11	7 077917	077917 oreochromis
10	22	36.1	11	7 077921	077921 pseudotroph
11	21	34.4	8	2 0989C2	Q9R9C2 borrelia bu
12	21	34.4	10	12 Q83978	Q83978 influenza a
13	19	31.1	10	6 Q9N1X1	Q9N1X1 equus caball
14	19	31.1	10	13 P82090	P82090 c. tritosa cit
15	18	29.5	10	13 073588	073588 gallus gall
16	18	29.5	11	11 P97755	P97755 rattus norv
17	18	29.5	11	13 P82092	P82092 littoria cit
18	17	27.9	7	2 Q34028	Q34028 pseudomonas
19	17	27.9	10	5 Q26093	Q26093 pisaster oc

20	17	27.9	11	2	068237	068237 borrelia bu
21	17	27.9	11	7	078121	078121 oreochromis
22	16	26.2	9	5	09TVF1	09TVF1 trypanosoma
23	16	26.2	11	6	09TRR7	09TRR7 oryctolagus
24	16	26.2	11	12	Q9DYW4	Q9DYW4 human immun
25	15	24.6	8	8	Q9T2W0	Q9T2W0 saccharomyc
26	15	24.6	8	4	Q9UGS4	Q9UGS4 homo sapien
27	15	24.6	9	12	Q70140	Q70140 human immun
28	15	24.6	11	13	P82088	P82088 littoria cit
29	15	24.6	11	13	P82089	P82089 littoria cit
30	14	23.0	8	2	087471	087471 haemophilus
31	14	23.0	8	3	P87225	P87225 saccharomyc
32	14	23.0	8	8	Q9T2V3	Q9T2V3 begonia for
33	14	23.0	9	4	Q9UNAO	Q9UNAO homo sapien
34	14	23.0	9	5	Q96417	Q96417 drosophila
35	14	23.0	9	6	Q9GK02	Q9GK02 capra hircu
36	14	23.0	10	2	P96421	P96421 neisseria g
37	14	23.0	10	2	Q9R5T2	Q9R5T2 acetobacter
38	14	23.0	10	11	Q63056	Q63056 rattus norv
39	14	23.0	11	2	P96319	P96319 desulfovibr.
40	14	23.0	11	2	Q9AIY6	Q9AIY6 candidatus
41	14	23.0	11	6	Q9TRX3	Q9TRX3 bos taurus
42	14	23.0	11	6	Q9TRX2	Q9TRX2 bos taurus
43	14	23.0	11	7	077883	077883 oreochromis
44	14	23.0	11	7	077886	077886 oreochromis
45	14	23.0	11	7	077912	077912 oreochromis

## ALIGNMENTS

RESULT 1

077899 PRELIMINARY; PRT; 11 AA.  
AC 077899;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE MHC CLASS II B LOCUS 14 (FRAGMENT).  
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;  
OC Cichlidae; Oreochromis.  
OX NCBI\_TaxID=8128;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98315113; PubMed=9649539;  
RA Malaga-Trillo E., McAndrew B., Vincsek V., Zaleska-Rutczynska Z.,  
RA Sueltmann H., Figueroa F., Klein J.;  
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
RT class II B loci."  
RL Genetics 149:1527-1547(1998).  
DR EMBL; AF050009; AAC41348.1; -.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 36.1%; Score 22; DB 7; Length 11;  
Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQDY 9  
: : : : :  
Db 3 MTYRLSRCD 11

RESULT 2  
077900 PRELIMINARY; PRT; 11 AA.  
AC 077900;

DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MHC CLASS II B LOCUS 14 (FRAGMENT).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98315113; PubMed=9649539;  
 RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,  
 RA Sueltmann H., Figueroa F., Klein J.;  
 RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
 class II B loci."  
 RL Genetics 149:1527-1547(1998).  
 DR EMBL; AF050010; AAC41349.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
 :||| :| :|  
 Db 3 MTYRLSRCD 11

RESULT 3

ID 077901 PRELIMINARY; PRT; 11 AA.  
 AC 077901;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MHC CLASS II B LOCUS 14 (FRAGMENT).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98315113; PubMed=9649539;  
 RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,  
 RA Sueltmann H., Figueroa F., Klein J.;  
 RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
 class II B loci."  
 RL Genetics 149:1527-1547(1998).  
 DR EMBL; AF050011; AAC41350.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
 :||| :| :|  
 Db 3 MTYRLSRCD 11

RESULT 4

077902  
 ID 077902 PRELIMINARY; PRT; 11 AA.  
 AC 077902;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MHC CLASS II B LOCUS 14 (FRAGMENT).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98315113; PubMed=9649539;  
 RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,  
 RA Sueltmann H., Figueroa F., Klein J.;  
 RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
 class II B loci."  
 RL Genetics 149:1527-1547(1998).  
 DR EMBL; AF050012; AAC41351.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
 :||| :| :|  
 Db 3 MTYRLSRCD 11

RESULT 5

ID 077903 PRELIMINARY; PRT; 11 AA.  
 AC 077903;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MHC CLASS II B LOCUS 14 (FRAGMENT).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidel;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98315113; PubMed=9649539;  
 RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,  
 RA Sueltmann H., Figueroa F., Klein J.;  
 RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
 class II B loci."  
 RL Genetics 149:1527-1547(1998).  
 DR EMBL; AF050013; AAC41352.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
 :||| :| :|  
 Db 3 MTYRLSRCD 11

```

RESULT 6
O77904 ID O77904 PRELIMINARY: PRT: 11 AA.
AC O77904;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS II B LOCUS 14 (FRAGMENT).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA Sueltmann H., Figueroa F., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT class II B loci.";
RL Genetics 149:1527-1547(1998).
DR EMBL; AF050014; AAC41353.1; -.
KW MHC.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

```

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Oy 1 LTYNTSQYD 9
Db 3 MTYRLSRCD 11

RESULT 7
O77905 ID O77905 PRELIMINARY: PRT: 11 AA.
AC O77905;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS II B LOCUS 14 (FRAGMENT).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA Sueltmann H., Figueroa F., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT class II B loci.";
RL Genetics 149:1527-1547(1998).
DR EMBL; AF050015; AAC41354.1; -.
KW MHC.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

```

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Oy 1 LTYNTSQYD 9
Db 3 MTYRLSRCD 11

RESULT 8
O77916 ID O77916 PRELIMINARY: PRT: 11 AA.
AC O77916;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS II B LOCUS 14 (FRAGMENT).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA Sueltmann H., Figueroa F., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT class II B loci.";
RL Genetics 149:1527-1547(1998).
DR EMBL; AF050029; AAC41368.1; -.
KW MHC.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

```

Query Match 36.1%; Score 22; DB 7; Length 11;  
 Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Oy 1 LTYNTSQYD 9
Db 3 MTYRLSRCD 11

RESULT 9
O77917 ID O77917 PRELIMINARY: PRT: 11 AA.
AC O77917;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS II B LOCUS 14 (FRAGMENT).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA Sueltmann H., Figueroa F., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT class II B loci.";
RL Genetics 149:1527-1547(1998).
DR EMBL; AF050030; AAC41369.1; -.
KW MHC.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

```

Query Match 36.1%; Score 22; DB 7; Length 11;

Best Local Similarity 44.4%; Pred. No. 8.5e+02; Mismatches 2; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
:||| :|  
Db 3 MTYRLSRCD 11

RESULT 10  
077921  
ID 077921 PRELIMINARY; PRT; 11 AA.  
AC 077921  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE MHC CLASS II B LOCUS 14 (FRAGMENT).  
OS Pseudotropheus sp. 'pseudotropheus tropheops complex'.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;  
OC Cichlidae; Pseudotropheus.  
OX NCBI\_TaxID=51796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98315113; PubMed=9649539;  
RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,  
RA Sueltmann H., Figueroa F., Klein J.;  
RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
RT Class II B loci".  
RL Genetics 149:1527-1547(1998).  
DR EMBL; AF050034; AAC41373.1; -.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 36.1%; Score 22; DB 7; Length 11;  
Best Local Similarity 44.4%; Pred. No. 8.5e+02;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
:||| :|  
Db 3 MTYRLSRCD 11

RESULT 11  
Q9R9C2  
ID Q9R9C2 PRELIMINARY; PRT; 8 AA.  
AC Q9R9C2  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE PLASMID CP32-3, POSSIBLE PARTITION PROTEINS, COMPLETE CDS (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OG Plasmid cp32-3.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=B31;  
RX MEDLINE=98361033; PubMed=9695920;  
RA Stevenson B., Castjens S., Rosa P.;  
RT "Evidence of past recombination events among the genes encoding the  
RT Erp antigens of Borrelia burgdorferi.";  
RL Microbiology 144:0-0(0).  
DR EMBL; AF022480; AAC35445.1; -.  
KW Plasmid.  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 985 MW; E8B41B41A735B446 CRC64;

Query Match 34.4%; Score 21; DB 2; Length 8;

Best Local Similarity 80.0%; Pred. No. 4.7e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTYNT 5  
:||| :|  
Db 4 LTYRT 8

RESULT 12  
Q83978  
ID Q83978 PRELIMINARY; PRT; 10 AA.  
AC Q83978  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)  
DE INFLUENZA A/UDORN/72 (H3N2), NUCLEOPROTEIN (SEG 5), 3' CDNA  
DE (FRAGMENT).  
OS Influenza A virus.  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group; Influenza A viruses.  
OX NCBI\_TaxID=11320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=A/UDORN/72 (H3N2);  
RX MEDLINE=83112211; PubMed=6296449;  
RA Lin B.C., Lai C.J.;  
RT "The influenza virus nucleoprotein synthesized from cloned DNA in a  
RT simian virus 40 vector is detected in the nucleus".  
RL J. Virol. 45:434-438(1983).  
DR EMBL; J02171; AAA43469.1; -.  
KW Nucleoprotein; Ribonucleoprotein.  
FT NON\_TER 1 1  
SQ SEQUENCE 10 AA; 1173 MW; 8787655B1B1DD44A CRC64;

Query Match 34.4%; Score 21; DB 12; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 NTSOYD 9  
:| :||  
Db 4 NAEYD 9

RESULT 13  
Q9N1X1  
ID Q9N1X1 PRELIMINARY; PRT; 10 AA.  
AC Q9N1X1  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE ALCOHOL DEHYDROGENASE 3 (FRAGMENT).  
GN ADH3.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20082971; PubMed=10613847;  
RA Caetano A.R., Shiue Y.L., Lyons L.A., O'Brien S.J., Laughlin T.F.,  
RA Bowling A.F., Murray J.D.;  
RT "A comparative gene map of the horse (Equus caballus).";  
RL Genome Res. 9:1239-1249(1999).  
DR EMBL; AF134056; AAF31299.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1144 MW; C4EA25676B02D6DD CRC64;

Query Match 31.1%; Score 19; DB 6; Length 10;  
Best Local Similarity 37.5%; Pred. No. 2.7e+03;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 YNTSQDYD 11  
 ||:| :  
 Db 1 NTAQVEPW 8

## RESULT 14

P82090  
 ID P82090 PRELIMINARY; PRT; 11 AA.  
 AC P82090;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE CAERULEIN 3.2/3.2Y4.  
 OS Litoria citropa (Australian blue mountains tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 OC Litoria.  
 OX NCBI\_TaxID=94770;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE-SKIN SECRETION;  
 RX MEDLINE=20057701; PubMed=10589099;  
 RA Wabnitz P.A., Bowie J.H., Tyler M.J.;  
 RT "Caerulein-like peptides from the skin glands of the Australian blue  
 mountains tree frog Litoria citropa. Part 1. Sequence determination  
 using electrospray mass spectrometry."; 2000.  
 RL Rapid Commun. Mass Spectrom. 13:2498-2502(1999).  
 CC -!- FUNCTION: HYPOTENSIVE NEUROPEPTIDE (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.  
 CC -!- MASS SPECTROMETRY: MW=1423; METHOD=ELECTROSPRAY.  
 CC -!- PTM: ISOFORM 3.2Y4 DIFFERS FROM ISOFORM 3.2 IN NOT BEING  
 CC SULFATED.  
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 KW Amphibian skin; Hypotensive agent; Amidation; Sulfatation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 4 4 SULFATATION.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1363 MW; 10DAB8867861A86B CRC64;

Query Match 31.1%; Score 19; DB 13; Length 11;  
 Best Local Similarity 42.9%; Pred. No. 2.9e+03;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 YNTSQDYD 9  
 ||:| :  
 Db 4 YGTGWFD 10

## RESULT 15

O73588  
 ID O73588 PRELIMINARY; PRT; 10 AA.  
 AC O73588;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE ENGRAILED-3 (FRAGMENT).  
 GN EN-3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WHITE LEGHORN; TISSUE-WHOLE EMBRYO;  
 RX MEDLINE=98141813; PubMed=9473273;  
 RA Peale F.V., Mason K., Hunter A.W., Bothwell M.;  
 RT "Multiplex display polymerase chain reaction amplifies and resolves  
 related sequences sharing a single moderately conserved domain."; 1998.  
 RL Anal. Biochem. 256:158-168(1998).  
 DR EMBL; U26148; AAC06186.1; -.

FT NON\_TER 1 1  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1118 MW; 73C0BE144735B72B CRC64;

Query Match 29.5%; Score 18; DB 13; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 4e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTYNTSQ 7  
 ||:| :  
 Db 4 LSLNESQ 10

Search completed: February 21, 2002, 17:23:06  
 Job time: 180 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2002, 17:16:50 ; Search time 23.72 seconds  
(without alignments)  
34.351 Million cell updates/sec

Title: US-08-753-851-16  
Perfect score: 61  
Sequence: 1 LTVNTSQYDPT 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 132282

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

```
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	11	15 AAR53486	CD44 peptide CD44-
2	28	45.9	11	6 AAP50473	Determinant site o
3	27	44.3	10	22 AAG85437	Saccharomyces cere
4	26.5	43.4	9	22 AAG88313	HER2/NEU DR superm
5	26	42.6	9	18 AAW24776	Human immunoglobul
6	26	42.6	9	19 AAW80146	Light chain CDR3 p
7	26	42.6	9	20 AAY40428	Amino acid sequenc
8	26	42.6	11	18 AAW15948	Interleukin-1 type
9	26	42.6	11	19 AAW68880	Peptide binding in
10	26	42.6	11	20 AAY10049	Interleukin-1 type
11	26	42.6	11	21 AAB17876	IL-1 antagonist pe

12	25	41.0	9	19 AAW59175	Mouse PAF receptor
13	25	41.0	10	20 AAY07198	HLA binding peptid
14	25	41.0	10	20 AAY07153	HLA binding peptid
15	25	41.0	10	22 AAG83833	Arbidopsis thalia
16	24	39.3	8	20 AAW97549	Antigenic site of
17	24	39.3	8	21 AAB06537	Claudin-2 cyclic c
18	24	39.3	9	22 AAB83062	Mealworm encapsula
19	24	39.3	10	22 AAG99126	Potato patatin pro
20	24	39.3	11	18 AAW20074	HGX7, a hexose oxi
21	23.5	38.5	11	20 AAY08919	S. peregrina cyste
22	23	37.7	6	16 AAR72626	Cladosporium herba
23	23	37.7	7	21 AAB23088	Peanut Ara h 2 mou
24	23	37.7	7	21 AAB33483	Mouse immunoglobul
25	23	37.7	7	21 AAB33534	Mouse immunoglobul
26	23	37.7	7	22 AAU04750	IgE binding epitop
27	23	37.7	7	22 AAU05075	Mouse IgE epitope
28	23	37.7	8	19 AAW63734	C. histolyticum CH
29	23	37.7	9	17 AAR93559	Dopamine receptor
30	23	37.7	9	20 AAY47044	Immunogenic peptid
31	23	37.7	9	22 AAG88701	HER2/NEU DR 3a mot
32	23	37.7	10	15 AAR50029	Leukocyte adhesion
33	23	37.7	10	17 AAR89619	Prostate specific
34	23	37.7	10	18 AAW33456	Oligopeptide 36 cl
35	23	37.7	10	18 AAW12686	Human urokinase-ty
36	23	37.7	10	20 AAY25037	Leukocyte adhesion
37	23	37.7	10	21 AAB27534	Ara h 2 linear iGE
38	23	37.7	10	21 AAY81797	Urokinase-type pla
39	23	37.7	10	22 AAU06346	Human Leukocyte An
40	23	37.7	10	22 AAG99122	Potato patatin pro
41	23	37.7	11	18 AAW19636	Non-covalently int
42	23	37.7	11	18 AAW15949	Interleukin-1 type
43	23	37.7	11	19 AAW68934	Peptide of the spe
44	23	37.7	11	20 AAY10050	Interleukin-1 type
45	23	37.7	11	20 AAW74276	HJ loop peptide K1

## ALIGNMENTS

RESULT 1  
AAR53486  
ID AAR53486 standard; peptide; 11 AA.  
XX  
AAR53486;  
XX  
01-DEC-1994 (first entry)  
XX  
CD44 peptide CD44-13.  
XX  
Cell adhesion molecule; CD44; antiinflammatory; rheumatoid;  
KW arthritis; tumor cell metastasis; autoimmune disease;  
KW immunosuppressive.  
XX  
Homo sapiens.  
XX  
WO9409811-A.  
XX  
11-MAY-1994.  
XX  
29-OCT-1993; 93WO-US10412.  
XX  
30-OCT-1992; 92US-0973339.  
PR  
(UYDU-) UNIV DUKE.  
XX  
Hale LP, Haynes BF, Liao H, Patton KL, Telen MJ;  
XX  
WPI; 1994-167121/20.  
XX  
Use of CD44 protein and new peptide derivs - for developing prods  
PT for inflammation, immune-mediated tissue damage and tumour cell  
PT metastasis  
XX

PS Claim 4; Page 14; 83pp; English.

XX The peptide can be used for treating inflammation and immune-mediated tissue damage such as occurs in the course of autoimmune diseases, e.g. rheumatoid arthritis. It can also be used for determining metastatic potential or to treat or prevent tumor cell metastasis. This peptide corresponds to AA 107-117 of the CD44 protein sequence.

XX Sequence 11 AA;

Query Match 100.0%; Score 61; DB 15; Length 11;  
Best Local Similarity 100.0%; Pred. No. 6.7e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTYNTSQYDTY 11  
| | | | | | | | | | |  
Db 1 ltyntsqtydy 11

RESULT 2

ARP50473  
ID AAP50473 standard; Protein; 11 AA.

XX AC AAP50473;

XX DT 14-JAN-1992 (first entry)

XX DE Determinant site of human somatomammotropin.

XX KW Somatomammotropin.

XX OS Synthetic.

XX PN US4493795-A.

XX PD 15-JAN-1985.

XX PF 17-OCT-1983; 83US-0542633.

XX PR 17-OCT-1983; 83US-0542633.

XX PA (SYNT ) SYNTEX (USA) INC.

XX PI Nestor JJ, Moffatt JG, Chan HW;

XX DR WPI; 1985-031644/05.

XX PT Synthetic antigenic polypeptide(s) - comprising amino acid sequences corresp. to beta-turn regions of proteins.

XX PS Claim 2; Page 14; 14pp; English.

XX The synthetic peptide sequence is useful as an antigen in the production of vaccines, antiviral agents, diagnostic reagents, etc., for the detection and therapy of infectious and immune diseases.

XX SQ Sequence 11 AA;

Query Match 45.9%; Score 28; DB 6; Length 11;  
Best Local Similarity 55.6%; Pred. No. 59;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 YNTSQYDTY 11

Db 1 ydtsdsdth 9

RESULT 3

AAG85437  
ID AAG85437 standard; Peptide; 10 AA.

XX AAG85437;

XX AC 11-SEP-2001 (first entry)

XX DE Saccharomyces cerevisiae peptide, SEQ ID NO: 386.

XX KW Saccharomyces cerevisiae; complementary peptide; peptide identification; drug discovery; drug design.

XX OS Saccharomyces cerevisiae.

XX PN WO200142276-A1.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB04773.

XX PR 13-DEC-1999; 99GB-0029471.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-367863/38.

XX PT Identifying complementary peptides by analysis of protein and nucleotide sequence databases, useful in drug design -

XX PS Example 3; Page 83; 488pp; English.

XX CC The invention relates to the identification of complementary peptides by analysis of protein and nucleotide sequence databases from higher eukaryotic genomes, excluding human and plants. The specific complementary peptides interact with their relevant target proteins encoded in the eukaryote genome. The peptides may be used as reagents and drugs for drug discovery and as lead ligands for drug design and development. The present sequence is a complementary peptide from Saccharomyces cerevisiae.

XX SQ Sequence 10 AA;

Query Match 44.3%; Score 27; DB 22; Length 10;  
Best Local Similarity 57.1%; Pred. No. 79;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 YNTSQYD 9

Db 1 fnnsryd 7

RESULT 4

AAG88313

ID AAG88313 standard; Peptide; 9 AA.

XX AC AAG88313;

XX DT 11-SEP-2001 (first entry)

XX DE HER2/NEU DR supermotif binding peptide core sequence #13.

XX KW Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell; immune response; vaccine; tumour; cancer; cytostatic; immunostimulant; tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200141787-A1.

XX PD 14-JUN-2001.

XX

PF 11-DEC-2000; 2000WO-US33591.  
 XX  
 PR 10-DEC-1999; 95US-0458299.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Cells E;  
 PI Keogh E;  
 XX  
 DR WPI; 2001-374995/39.  
 XX  
 XX  
 PT An isolated prepared HER2/neu epitope useful in a vaccine for inducing  
 PT cellular immune responses for the prevention and treatment of cancer -  
 XX  
 XX Disclosure; Page 166; 199pp; English.  
 XX  
 CC The present invention describes isolated prepared HER2/neu epitopes (I).  
 CC Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is  
 CC culture in vitro and binds to a complex of an epitope (I), bound to a  
 CC human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)  
 CC and a second epitope and the peptide is less than 50 contiguous amino  
 CC acids that have 100% identity with a native peptide sequence of HER2/neu;  
 CC (3) a vaccine composition (III) comprising (II) and a pharmaceutical  
 CC excipient; (4) an isolated nucleic acid encoding a peptide comprising  
 CC (I); and (5) an isolated nucleic acid encoding (II). (I) has cytostatic  
 CC and immunostimulant activities, and can be used in vaccines. (I), (II)  
 CC and (III) are useful for inducing cellular immune responses for the  
 CC prevention and treatment of cancer. (I) and (II) are useful for  
 CC monitoring or evaluating an immune response to a tumour-associated  
 CC antigen when incubated with a T lymphocyte sample from a patient and  
 CC detecting the presence of bound T lymphocyte to (I) or (II). Epitope  
 CC based vaccines mean that immunosuppressive epitopes that may be present  
 CC in whole antigens may be avoided. Selected epitopes may be combined to  
 CC enhance immunogenicity. The possible pathological side effects caused by  
 CC infectious agents or whole protein antigen is eliminated. The vaccine  
 CC provides the ability to direct and focus an immune response to multiple  
 CC selected antigens from the same pathogen. Epitope-based anti-tumour  
 CC vaccines provides the opportunity to combine epitopes derived from  
 CC multiple tumour-associated molecules addressing the problem of tumour-  
 CC tumour variability and reducing the likelihood of tumour escape due to  
 CC antigen loss. AAG8266 to AAG89121 represent amino acid sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 43.4%; Score 26.5; DB 22; Length 9;  
 Best Local Similarity 54.5%; Pred. No. 4.3e+05;  
 Matches 6; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 1 LRYNTSQVDY 11  
 :|||||  
 Db 2 vtyn---dtf 9

RESULT 5  
 AAW24776  
 ID AAW24776 standard; peptide; 9 AA.  
 XX  
 AC AAW24776;  
 XX  
 DT 03-DEC-1997 (first entry)  
 XX  
 DE Human immunoglobulin light chain CDR3 region peptide.  
 XX  
 KW Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;  
 KW transgenic; mouse; CD4; antibody; autoimmune; inflammatory;  
 KW transplant rejection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9713852-A1.  
 XX

Query Match 43.4%; Score 26.5; DB 22; Length 9;  
 Best Local Similarity 54.5%; Pred. No. 4.3e+05;  
 Matches 6; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 1 LRYNTSQVDY 11  
 :|||||  
 Db 2 vtyn---dtf 9

RESULT 5  
 AAW24776  
 ID AAW24776 standard; peptide; 9 AA.  
 XX  
 AC AAW24776;  
 XX  
 DT 03-DEC-1997 (first entry)  
 XX  
 DE Human immunoglobulin light chain CDR3 region peptide.  
 XX  
 KW Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;  
 KW transgenic; mouse; CD4; antibody; autoimmune; inflammatory;  
 KW transplant rejection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9713852-A1.  
 XX

PD 17-APR-1997.  
 XX  
 PF 10-OCT-1996; 96WO-US16433.  
 XX  
 PR 10-OCT-1995; 95US-0544404.  
 XX  
 PA (GENP-) GENPHARM INT INC.  
 XX  
 PI Kay RM; Lonberg N;  
 XX  
 DR WPI; 1997-235888/21.  
 XX  
 PT Novel anti-CD4 antibody produced by transgenic mice - used in the  
 PT treatment of auto-immune disease etc.  
 XX  
 XX Claim 60; Page 287; 396pp; English.  
 PS  
 XX A novel composition has been developed which comprises an immunoglobulin  
 CC (Ig) having an affinity constant (Ka) of at least 2 multiply  
 CC 100000000 M-1 for binding to a predetermined human antigen. The  
 CC present sequence represents a human light chain CDR3 region peptide  
 CC which forms part of an immunoglobulin comprising a VKL15 segment,  
 CC a Jk2 segment, and the present light chain CDR3 region. The  
 CC anti-CD4 antibodies may be used in therapeutic and diagnostic  
 CC applications, especially for the treatment of human diseases. These  
 CC antibodies reduce activity of CD4 cells and reduce undesirable  
 CC autoimmune reactions, inflammatory response and transplant rejection.  
 CC Transgenic animals are capable of producing heterologous antibodies  
 CC of multiple isotypes by undergoing isotype switching. These animals  
 CC produce a first Ig type that is necessary for antigen-stimulated B-cell  
 CC maturation and can switch to encode and produce one or more subsequent  
 CC heterologous isotypes.  
 XX  
 SQ Sequence 9 AA;

Query Match 42.6%; Score 26; DB 18; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 QYDXY 11  
 :|||:  
 Db 2 qydsy 6

RESULT 6  
 AAW80146  
 ID AAW80146 standard; peptide; 9 AA.  
 XX  
 AC AAW80146;  
 XX  
 DT 17-DEC-1998 (first entry)  
 XX  
 DE Light chain CDR3 peptide from clone 4D1.  
 XX  
 KW Complementarity determining region; CDR3; transgenic non-human animal;  
 KW human heterologous antibody; heavy chain; light chain transgene;  
 KW neutrophil efflux; reperfusion injury; CD4 binding; autoimmune reaction;  
 KW inflammatory response; transplant rejection; acid induced lung injury;  
 KW acute adult respiratory distress syndrome; ARDS; vasculitis;  
 KW septic shock; allergic reaction; asthma; cystic fibrosis.  
 XX  
 OS Chimeric - Mus sp.  
 OS Chimeric - Homo sapiens.  
 XX  
 PN WO9824884-A1.  
 XX  
 PD 11-JUN-1998.  
 XX  
 PF 01-DEC-1997; 97WO-US21803.  
 XX  
 PR 02-DEC-1996; 96US-0758417.  
 XX

PA (GENP-) GENPHARM INT.

XX Kay RM, Lonberg N;

XX WPI; 1998-333306/29.

XX Hybridoma producing antibody specific for interleukin-8 - used to  
PT prevent efflux of neutrophils from vasculature, and treat  
PT reperfusion injury

XX Example 41; Page 306; 452pp; English.

XX The present sequence represents a light chain complementarity  
CC determining region (CDR) 3 peptide. It is part of the antibody of  
CC the invention. The specification describes transgenic non-human  
CC animals, especially a mouse, which are capable of producing a human  
CC heterologous antibodies of multiple isotypes by undergoing isotype  
CC switching. The transgenic animals have human heavy and light chain  
CC transgenes. The transgenes are capable of functionally rearranging a  
CC heterologous diversity (D) gene in a variable-diversity-junction (V-D-J)  
CC recombination. The transgenes include a heavy chain transgene comprising  
CC at least one V, D and J gene segment, and one constant region gene  
CC segment. The immunoglobulin (Ig) light chain transgene comprises at least  
CC one V and J gene segment and one constant region gene segment. The gene  
CC segments are heterologous to the transgenic animal. The antibody can be  
CC used to prevent efflux of neutrophils from vasculature. It can also be  
CC used to treat reperfusion injury. CD4 binding antibodies are used to  
CC reduce undesirable autoimmune reactions, inflammatory responses and  
CC rejection of transplanted organs. The anti-IL-8 antibodies can reduce  
CC tissue damage and prolong survival in animal models of acute adult  
CC respiratory distress syndrome (ARDS) and acid induced lung injury. The  
CC anti-IL-8 antibodies can also be used for the treatment of vasculitis,  
CC septic shock, allergic reactions (e.g. asthma) and cystic fibrosis.  
CC note: mouse and human parts of the sequence are not defined in the  
CC specification.

XX Sequence 9 AA;

Query Match 42.6%; Score 26; DB 19; Length 9;

Best Local Similarity 80.0%; Pred. NO. 4.3e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 QYDTY 11

Db 2 qydsy 6

RESULT 7

AA40428

ID AAY40428 standard; Peptide; 9 AA.

XX AC AAY40428;

XX DT 24-NOV-1999 (first entry)

DE Amino acid sequence of the VDJ joint of a heterologous antibody.

XX Transgenic animal; heterologous antibody; hybridoma; B cell;

KW transgenic mouse; human heavy chain transgene; digoxin;

KW human light chain transgene; immortalized cell; immunoglobulin;

KW Shinga-like toxin; autoimmune disease; cancer; infectious disease;

KW transplant rejection; blood disorder; coagulation disorder.

XX Synthetic.

XX WO945962-A1.

XX 16-SEP-1999.

XX PF 12-MAR-1999; 99WO-US05535.

XX PR 13-MAR-1998; 98US-0042353.

XX (GENP-) GENPHARM INT INC.

XX Lonberg N, Fishwild DM, Ball WJ;

XX WPI; 1999-551219/46.

XX Novel transgenic non-human animals used to produce heterologous

XX antibodies

XX Example 41; Page 307; 484pp; English.

XX The specification describes transgenic animals that are capable of  
CC producing a heterologous antibody. The antibodies are isolated from a  
CC hybridoma, comprising B cells, that is obtained from a transgenic mouse  
CC having a genome comprising a human heavy chain transgene and a human  
CC light chain transgene. The B cells are fused to immortalized cells  
CC suitable for generating a hybridoma, which produces a detectable  
CC amount of an immunoglobulin that specifically binds digoxin or  
CC Shinga-like toxin. B cells from transgenic animals can be used to  
CC generate hybridomas expressing monoclonal high affinity human sequence  
CC antibodies. Antibodies produced from the transgenic animals of the  
CC invention can be used to treat human diseases, e.g. autoimmune  
CC diseases, cancer, infectious disease, transplant rejection, blood  
CC disorders, such as coagulation disorders and other diseases. The  
CC present sequence represents a VDJ joint of a heterologous antibody  
CC of the invention.

XX Sequence 9 AA;

Query Match 42.6%; Score 26; DB 20; Length 9;

Best Local Similarity 80.0%; Pred. NO. 4.3e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 QYDTY 11

Db 2 qydsy 6

RESULT 8

AAW15948

ID AAW15948 standard; peptide; 11 AA.

XX AC AAW15948;

XX DT 19-AUG-1997 (first entry)

DE Interleukin-1 type I receptor binding peptide.

XX IL-1RtI; agonist; antagonist.

XX Synthetic.

XX WO9639165-A1.

XX PD 12-DEC-1996.

XX PF 05-JUN-1996; 96WO-US09835.

XX PR 05-JUN-1995; 95US-0464538.

XX {AFFY-) AFFYMAX TECHNOLOGIES NV.

XX Baldwin D, Barrett RW, Bovy PR, Pharanipragada R;

XX Jacobs JW, Leahy EM, Pottorf RS, Tomlinson RC, Yanofsky SD;

XX WPI; 1997-042846/04.

XX Interleukin-1 type I receptor inhibitor peptide(s) and compounds  
PT used to antagonise the activity of IL-1, for treatment of e.g. AIDS,  
PT rheumatoid arthritis, chronic hepatitis B, etc.



CC These peptides are useful as agonists/antagonists for the study of  
 CC IL-1 mediated activities (e.g. as labels and probes), for the  
 CC identification of new IL-1 receptor blockers, and for the  
 CC identification, diagnosis and treatment/prevention of diseases with  
 CC an inappropriate production or response to IL-1, e.g. rheumatoid  
 CC arthritis, osteoarthritis, psoriasis, inflammatory bowel disease,  
 CC encephalitis, glomerulonephritis, and respiratory distress syndrome.  
 CC The present sequence is a specifically claimed example of the above  
 CC peptides.

XX Sequence 11 AA;

Query Match 42.6%; Score 26; DB 20; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NTSQYDTY 11  
 I: I I I I:  
 Db 2 nsswydsf 9

## RESULT 11

AAAB17876  
 ID AAB17876 standard; Peptide: 11 AA.

XX AAB17876;

DT 31-OCT-2000 (first entry)

DE IL-1 antagonist peptide sequence SEQ ID NO:980.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
 KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase;  
 KW asthma; thrombosis; pharmaceutical.

XX Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US25044.

XX 23-OCT-1998; 98US-0105371.

XX 22-OCT-1999; 99US-0428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheatham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and  
 PT pharmacologically active peptides, useful for treating cancer and  
 PT autoimmune diseases -

XX Claim 10; Page 538; 608pp; English.

XX The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,  
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4  
 CC where P1, P2, P3, and P4 = are each independently sequences of  
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each  
 CC independently linkers; and a, b, c, d, e, and f = are each independently  
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can  
 CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive

CC activities. DNAs, vectors and host cells from the present invention can  
 CC be used for producing pharmaceutical compositions. The compositions are  
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer  
 CC half-life or incorporate functions such as Fc receptor binding, protein  
 CC A binding, complement fixation, and possibly placental transfer. AAA69443  
 CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid  
 CC sequences used in the exemplification of the present invention.

XX Sequence 11 AA;

Query Match 42.6%; Score 26; DB 21; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NTSQYDTY 11  
 I: I I I I:  
 Db 2 nsswydsf 9

## RESULT 12

AAW59175  
 ID AAW59175 standard; Protein; 9 AA.

XX AAW59175;

DT 13-AUG-1998 (first entry)

DE Mouse PAF receptor antibody light chain hypervariable region CDR3.

XX Immunoglobulin light chain; hypervariable region; PAF; CDR; antibody;  
 KW platelet activating factor receptor; murine; chimeric; anti-PAF;  
 KW complementarity determining region.

XX Mus sp.

XX JPI0136981-A.

XX 26-MAY-1998.

XX 12-NOV-1996; 96JP-0317047.

XX 12-NOV-1996; 96JP-0317047.

XX (SUME ) SUMITOMO ELECTRIC IND CO.

XX WPI; 1998-355043/31.

XX Polypeptide comprising variable region of antibody recognising PAF  
 PT receptor - useful for the preparation of chimeric antibody

XX Claim 2; Page 10; 14pp; Japanese.

XX This sequence represents the complementarity determining region (CDR)  
 CC 3 of the light chain hypervariable region of an antibody recognising  
 CC the platelet activating factor (PAF) receptor. The variable region of  
 CC mouse anti-PAF receptor monoclonal antibody is useful for the  
 CC preparation of chimeric antibodies.

XX Sequence 9 AA;

Query Match 41.0%; Score 25; DB 19; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 QYDTY 11  
 I: I I I I:  
 Db 2 qyntq 6

## RESULT 13

AAY07198  
 ID AAY07198 standard; Peptide; 10 AA.  
 XX  
 AC AAY07198;  
 XX  
 DT 02-JUL-1999 (first entry)  
 DE HLA binding peptide.  
 XX  
 DE HLA binding peptide.  
 KW Cancer associated antigen; diagnosis; research; treatment; human;  
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
 KW prostate cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9904265-A2.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PF 15-JUL-1998; 98WO-US14679.  
 XX  
 PR 22-JUN-1998; 98US-0102322.  
 PR 17-JUL-1997; 97US-0896164.  
 PR 10-OCT-1997; 97US-0061599.  
 PR 10-OCT-1997; 97US-0061765.  
 PR 10-OCT-1997; 97US-0948705.  
 PR 11-OCT-1997; 97GB-0021697.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;  
 PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;  
 PI Tureci O;  
 XX  
 DR WPI; 1999-132448/11.  
 XX  
 PT New isolated cancer associated nucleic acids and polypeptides -  
 PT isolated using sera from cancer patients, used to develop products  
 PT for the diagnosis, monitoring or treatment of cancers  
 XX  
 PS Example 16; Page 767; 787pp; English.  
 XX  
 CC The invention relates to a method for diagnosing a disorder characterised  
 CC by expression of a human cancer associated antigen precursor coded for by  
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a  
 CC biological sample isolated from a subject with an agent that specifically  
 CC binds to the NAM, an expression product or a fragment of an expression  
 CC product complexed with an HLA molecule; and (b) determining the  
 CC interaction between the agent and the NAM or the expression product as a  
 CC determination of the disorder. The products and methods can be used in  
 CC the diagnosis, monitoring, research, or treatment of conditions  
 CC characterised by the expression of various cancer associated antigens.  
 CC The invention provides nucleic acid sequences and encoded polypeptides  
 CC which are cancer associated antigen precursors expressed in human breast  
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
 CC lung cancer.  
 XX  
 SQ Sequence 10 AA;  
  
 Query Match 41.0%; Score 25; DB 20; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 3 YNTSQYD 9  
 || : ||  
 Db 3 ynxtygd 9  
  
 RESULT 14  
 AAY07153  
 ID AAY07153 standard; Peptide; 10 AA.  
 XX

AAY07153;  
 AC  
 XX  
 DT 02-JUL-1999 (first entry)  
 XX  
 DE HLA binding peptide.  
 XX  
 DE HLA binding peptide.  
 KW Cancer associated antigen; diagnosis; research; treatment; human;  
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
 KW prostate cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9904265-A2.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PF 15-JUL-1998; 98WO-US14679.  
 XX  
 PR 22-JUN-1998; 98US-0102322.  
 PR 17-JUL-1997; 97US-0896164.  
 PR 10-OCT-1997; 97US-0061599.  
 PR 10-OCT-1997; 97US-0061765.  
 PR 10-OCT-1997; 97US-0948705.  
 PR 11-OCT-1997; 97GB-0021697.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;  
 PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;  
 PI Tureci O;  
 XX  
 DR WPI; 1999-132448/11.  
 XX  
 PT New isolated cancer associated nucleic acids and polypeptides -  
 PT isolated using sera from cancer patients, used to develop products  
 PT for the diagnosis, monitoring or treatment of cancers  
 XX  
 PS Example 16; Page 760; 787pp; English.  
 XX  
 CC The invention relates to a method for diagnosing a disorder characterised  
 CC by expression of a human cancer associated antigen precursor coded for by  
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a  
 CC biological sample isolated from a subject with an agent that specifically  
 CC binds to the NAM, an expression product or a fragment of an expression  
 CC product complexed with an HLA molecule; and (b) determining the  
 CC interaction between the agent and the NAM or the expression product as a  
 CC determination of the disorder. The products and methods can be used in  
 CC the diagnosis, monitoring, research, or treatment of conditions  
 CC characterised by the expression of various cancer associated antigens.  
 CC The invention provides nucleic acid sequences and encoded polypeptides  
 CC which are cancer associated antigen precursors expressed in human breast  
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
 CC lung cancer.  
 XX  
 SQ Sequence 10 AA;  
  
 Query Match 41.0%; Score 25; DB 20; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 3 YNTSQYD 9  
 || : ||  
 Db 3 ynxtygd 9  
  
 RESULT 15  
 AAG83833  
 ID AAG83833 standard; Peptide; 10 AA.  
 XX  
 AC AAG83833;  
 XX  
 DT 11-SEP-2001 (first entry)

XX Arabidopsis thaliana peptide ligand #473.  
 XX Plant; peptide pesticide; peptide herbicide; agricultural research.  
 KW Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX WO200142279-A2.  
 PN 14-JUN-2001.  
 PD 13-DEC-2000; 2000WO-GE04781.  
 PF 13-DEC-1999; 99GB-0029469.  
 PR (PROT-) PROTEOM LTD.  
 PA Roberts GW, Heal JR;  
 PI WPI; 2001-381629/40.  
 XX A set of peptide ligands for agricultural research and development of  
 PT therapeutic agents comprise specific complementary peptides to proteins  
 PT encoded by genes of plant genomes.  
 XX Example 4; Page 97; 201pp; English.  
 PS The present invention relates to a set of peptide ligands consisting of  
 XX specific complementary peptides to proteins encoded by genes of plant  
 CC genomes. The present sequence is one such peptide from Arabidopsis  
 CC thaliana. The peptides of the present invention are useful in an assay to  
 CC identify a peptide, especially a peptide pesticide or herbicide. The  
 CC peptides are also useful for tools for agricultural research and  
 CC development.  
 XX Sequence 10 AA;  
 SQ

Query Match 41.0%; Score 25; DB 22; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 LTYNTSQYDF 10  
 |||.|:  
 Db 1 ltysttdst 10

Search completed: February 21, 2002, 17:20:02  
 Job time: 192 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: February 21, 2002, 17:19:20 ; Search time 90.49 Seconds  
(without alignments)  
33.752 Million cell updates/sec

Title: US-08-753-851-16

Perfect score: 61

Sequence: 1 LTNWTSQYDIY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 279397

Minimum DB seq length: 0

Maximum DB seq length: 11

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA.Main.\*

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	61	100.0	11	3	US-07-973-339-14
2	61	100.0	11	4	US-08-047-068-16
3	61	100.0	11	5	US-08-143-311A-14
4	61	100.0	11	11	US-08-753-851-16
5	61	100.0	11	13	US-08-959-784-14
6	29	47.5	11	18	US-09-458-299-800
7	29	47.5	11	18	US-09-458-299-1422
8	29	47.5	11	18	US-09-458-299-1998
9	28	45.9	10	18	US-09-458-299-1399

10	27	44.3	9	21	US-09-793-451-152	Sequence 152, App
11	27	44.3	9	21	US-09-793-451-248	Sequence 248, App
12	27	44.3	9	21	US-09-793-451-564	Sequence 564, App
13	27	44.3	10	21	US-09-793-451-602	Sequence 602, App
14	27	44.3	10	21	US-09-793-451-607	Sequence 607, App
15	26.5	43.4	8	18	US-09-458-299-2	Sequence 2, Appli
16	26.5	43.4	8	18	US-09-458-299-1421	Sequence 1421, Ap
17	26.5	43.4	8	18	US-09-458-299-1997	Sequence 1997, Ap
18	26.5	43.4	8	18	US-09-458-299-3140	Sequence 3140, Ap
19	26.5	43.4	9	18	US-09-458-299-53	Sequence 53, Appl
20	26.5	43.4	9	18	US-09-458-299-1253	Sequence 1253, Ap
21	26.5	43.4	9	18	US-09-458-299-2191	Sequence 2191, Ap
22	26.5	43.4	9	18	US-09-458-299-2863	Sequence 2863, Ap
23	26.5	43.4	9	18	US-09-458-299-3966	Sequence 3966, Ap
24	26.5	43.4	10	18	US-09-458-299-77	Sequence 77, Appl
25	26.5	43.4	10	18	US-09-458-299-1046	Sequence 1046, Ap
26	26.5	43.4	10	18	US-09-458-299-2008	Sequence 2008, Ap
27	26.5	43.4	10	18	US-09-458-299-2482	Sequence 2482, Ap
28	26.5	43.4	11	18	US-09-458-299-1913	Sequence 1913, Ap
29	26.5	43.4	11	18	US-09-458-299-2900	Sequence 2900, Ap
30	26	42.6	9	11	US-08-728-463-216	Sequence 216, App
31	26	42.6	9	11	US-08-728-463B-216	Sequence 980, App
32	26	42.6	11	1	PCT-US99-25044-980	Sequence 280, App
33	26	42.6	11	8	US-08-458-593-280	Sequence 280, App
34	26	42.6	11	8	US-08-463-076D-280	Sequence 280, App
35	26	42.6	11	8	US-08-465-268B-280	Sequence 280, App
36	26	42.6	11	18	US-09-428-082-980	Sequence 980, App
37	26	42.6	11	21	US-09-751-181-28	Sequence 28, Appl
38	25	41.0	10	19	US-09-572-270A-473	Sequence 473, App
39	24	39.3	8	15	US-09-185-908-126	Sequence 126, App
40	24	39.3	8	16	US-09-282-029-126	Sequence 126, App
41	24	39.3	8	18	US-09-434-355-126	Sequence 126, App
42	24	39.3	8	18	US-09-434-355A-126	Sequence 126, App
43	24	39.3	8	18	US-09-458-299-517	Sequence 517, App
44	24	39.3	8	18	US-09-458-299-561	Sequence 561, App
45	24	39.3	9	18	US-09-458-299-150	Sequence 150, App

#### ALIGNMENTS

RESULT 1  
US-07-973-339-14  
; Sequence 14, Application US/07973339  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/973,339  
; FILING DATE: 19921030  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEST JR., WILLIAM K.  
; REGISTRATION NUMBER: 22057  
; REFERENCE/DOCKET NUMBER: 95459  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3000  
; TELEFAX: 202-822-0944

; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-973-339-14

Query Match 100.0%; Score 61; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTYNTSQYDTY 11  
| | | | | | | | | | |  
Db 1 LTYNTSQYDTY 11

RESULT 2  
US-08-047-068-16  
; Sequence 16, Application US/08047068  
; GENERAL INFORMATION:  
; APPLICANT: WEINBERG, J. BRICE  
; APPLICANT: HAYNES, BARTON F.  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/047,068  
; FILING DATE: 16-APR-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/945,581  
; FILING DATE: 16-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/682,518  
; FILING DATE: 09-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/669,730  
; FILING DATE: 15-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-21  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; TELEX: 200797 NIXN UR  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-047-068-16

Query Match 100.0%; Score 61; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTYNTSQYDTY 11  
| | | | | | | | | | |  
Db 1 LTYNTSQYDTY 11

RESULT 3  
US-08-143-311A-14  
; Sequence 14, Application US/08143311A  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311A  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; TELEX: 200797 NIXN UR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-143-311A-14

Query Match 100.0%; Score 61; DB 5; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTYNTSQYDTY 11  
| | | | | | | | | | |  
Db 1 LTYNTSQYDTY 11

RESULT 4  
US-08-753-851-16  
; Sequence 16, Application US/08753851  
; GENERAL INFORMATION:  
; APPLICANT: WEINBERG, J. BRICE  
; APPLICANT: HAYNES, BARTON F.  
; TITLE OF INVENTION: AN ADHESION MOLECULE

NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/753,851  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/047,068  
FILING DATE: 16-APR-1993  
APPLICATION NUMBER: US 07/945,581  
FILING DATE: 16-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/682,518  
FILING DATE: 09-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/669,730  
FILING DATE: 15-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-753-851-16

Query Match 100.0%; Score 61; DB 11; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTYNTSQYDYY 11  
Db 1 LTYNTSQYDYY 11

RESULT 5  
US-08-959-784-14  
Sequence 14, Application US/08959784  
GENERAL INFORMATION:  
APPLICANT: HAYNES, BARTON F.  
APPLICANT: HALE, LAURA P.  
APPLICANT: PATTON, KAREN L.  
APPLICANT: TELEN, MARILYN J.  
APPLICANT: LIAO, HUA-XIN  
TITLE OF INVENTION: AN ADHESION MOLECULE  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/959,784  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/143,311  
FILING DATE: 29-OCT-1993  
APPLICATION NUMBER: 07/973,339  
FILING DATE: 30-OCT-1992  
APPLICATION NUMBER: 07/669,730  
FILING DATE: 15-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 1579-50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-959-784-14

Query Match 100.0%; Score 61; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTYNTSQYDYY 11  
Db 1 LTYNTSQYDYY 11

RESULT 6  
US-09-458-299-800  
Sequence 800, Application US/09458299  
GENERAL INFORMATION:  
APPLICANT: Epimmune  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO  
TITLE OF INVENTION: HER2/neu USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
FILE REFERENCE: 018623-0148000S  
CURRENT APPLICATION NUMBER: US/09/458,299  
CURRENT FILING DATE: 2001-09-17  
PRIOR APPLICATION NUMBER: US 09/189,702  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: US 08/205,713  
PRIOR FILING DATE: 1994-03-04  
PRIOR APPLICATION NUMBER: US 08/159,184  
PRIOR FILING DATE: 1993-11-29  
PRIOR APPLICATION NUMBER: US 08/073,205  
PRIOR FILING DATE: 1993-06-04  
PRIOR APPLICATION NUMBER: US 08/027,146  
PRIOR FILING DATE: 1993-03-05  
NUMBER OF SEQ ID NOS: 4570  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 800  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-09-458-299-800

Query Match 47.5%; Score 29; DB 18; Length 11;  
Best Local Similarity 40.0%; Pred. No. 2.5e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTYNTSQYDT 10  
:|||| :  
Db 1 VTYNTDTFES 10

RESULT 7  
US-09-458-299-1422  
; Sequence 1422, Application US/09458299  
; GENERAL INFORMATION:  
; APPLICANT: Epimmune  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Keogh, Elissa  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO  
; FILE OF INVENTION: HER2/neu USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 018623-014800US  
; CURRENT APPLICATION NUMBER: US/09/458,299  
; CURRENT FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/159,184  
; PRIOR FILING DATE: 1993-11-29  
; PRIOR APPLICATION NUMBER: US 08/073,205  
; PRIOR FILING DATE: 1993-06-04  
; PRIOR APPLICATION NUMBER: US 08/027,146  
; NUMBER OF SEQ ID NOS: 4570  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1422  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-09-458-299-1422

Query Match 47.5%; Score 29; DB 18; Length 11;  
Best Local Similarity 40.0%; Pred. No. 2.5e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTYNTSQYDT 10  
:|||| :  
Db 1 VTYNTDTFES 10

RESULT 8  
US-09-458-299-1998  
; Sequence 1998, Application US/09458299  
; GENERAL INFORMATION:  
; APPLICANT: Epimmune  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Keogh, Elissa

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO  
; FILE OF INVENTION: HER2/neu USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 018623-014800US  
; CURRENT APPLICATION NUMBER: US/09/458,299  
; CURRENT FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/159,184  
; PRIOR FILING DATE: 1993-11-29  
; PRIOR APPLICATION NUMBER: US 08/073,205  
; PRIOR FILING DATE: 1993-06-04  
; PRIOR APPLICATION NUMBER: US 08/027,146  
; NUMBER OF SEQ ID NOS: 4570  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1998  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-09-458-299-1998

Query Match 47.5%; Score 29; DB 18; Length 11;  
Best Local Similarity 40.0%; Pred. No. 2.5e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTYNTSQYDT 10  
:|||| :  
Db 1 VTYNTDTFES 10

RESULT 9  
US-09-458-299-1399  
; Sequence 1399, Application US/09458299  
; GENERAL INFORMATION:  
; APPLICANT: Epimmune  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Keogh, Elissa  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO  
; FILE OF INVENTION: HER2/neu USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 018623-014800US  
; CURRENT APPLICATION NUMBER: US/09/458,299  
; CURRENT FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/159,184  
; PRIOR FILING DATE: 1993-11-29  
; PRIOR APPLICATION NUMBER: US 08/073,205  
; PRIOR FILING DATE: 1993-06-04  
; PRIOR APPLICATION NUMBER: US 08/027,146  
; NUMBER OF SEQ ID NOS: 4570  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1399  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-09-458-299-1399

Query Match 45.9%; Score 28; DB 18; Length 10;  
 Best Local Similarity 44.4%; Pred. No. 3.4e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 2 TYNTSOYDT 10  
 . . . . .  
 Db 1 TYNTDTFES 9

RESULT 10  
 US-09-793-451-152  
 ; Sequence 152, Application US/09793451  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Gazelle S. Rastegar  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Mary Paris  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
 ; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
 ; FILE REFERENCE: 129.2USU2  
 ; CURRENT APPLICATION NUMBER: US/09/793,451  
 ; CURRENT FILING DATE: 2001-02-26  
 ; PRIOR APPLICATION NUMBER: 60/184,558  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/218,856  
 ; PRIOR FILING DATE: 2000-07-13  
 ; NUMBER OF SEQ ID NOS: 752  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 152  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-793-451-152

Query Match 44.3%; Score 27; DB 21; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 5 TSQYDT 10  
 . . . . .  
 Db 4 THQYDT 9

RESULT 11  
 US-09-793-451-248  
 ; Sequence 248, Application US/09793451  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Gazelle S. Rastegar  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Mary Paris  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
 ; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
 ; FILE REFERENCE: 129.2USU2  
 ; CURRENT APPLICATION NUMBER: US/09/793,451  
 ; CURRENT FILING DATE: 2001-02-26  
 ; PRIOR APPLICATION NUMBER: 60/184,558  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/218,856  
 ; PRIOR FILING DATE: 2000-07-13  
 ; NUMBER OF SEQ ID NOS: 752  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 248  
 ; LENGTH: 9

; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-793-451-248

Query Match 44.3%; Score 27; DB 21; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 5 TSQYDT 10  
 . . . . .  
 Db 4 THQYDT 9

RESULT 12  
 US-09-793-451-564  
 ; Sequence 564, Application US/09793451  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Gazelle S. Rastegar  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Mary Paris  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
 ; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
 ; FILE REFERENCE: 129.2USU2  
 ; CURRENT APPLICATION NUMBER: US/09/793,451  
 ; CURRENT FILING DATE: 2001-02-26  
 ; PRIOR APPLICATION NUMBER: 60/184,558  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/218,856  
 ; PRIOR FILING DATE: 2000-07-13  
 ; NUMBER OF SEQ ID NOS: 752  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 564  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-793-451-564

Query Match 44.3%; Score 27; DB 21; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 5 TSQYDT 10  
 . . . . .  
 Db 2 THQYDT 7

RESULT 13  
 US-09-793-451-602  
 ; Sequence 602, Application US/09793451  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Gazelle S. Rastegar  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Mary Paris  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
 ; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
 ; FILE REFERENCE: 129.2USU2  
 ; CURRENT APPLICATION NUMBER: US/09/793,451  
 ; CURRENT FILING DATE: 2001-02-26  
 ; PRIOR APPLICATION NUMBER: 60/184,558  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/218,856

; PRIOR FILING DATE: 2000-07-13  
; NUMBER OF SEQ ID NOS: 752  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 602  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-793-451-602

Query Match 44.3%; Score 27; DB 21; Length 10;  
Best Local Similarity 83.3%; Pred. No. 5.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TSQYDT 10  
| | | | |  
Db 3 THQYDT 8

## RESULT 14

US-09-793-451-607  
; Sequence 607, Application US/09793451  
; GENERAL INFORMATION:  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Gazelle S. Rastegar  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Pia M. Challita-Eld  
; APPLICANT: Mary Faris  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
; FILE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129 2USU2  
; CURRENT APPLICATION NUMBER: US/09/793,451  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/184,558  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/218,856  
; PRIOR FILING DATE: 2000-07-13  
; NUMBER OF SEQ ID NOS: 752  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 607  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-793-451-607

Query Match 44.3%; Score 27; DB 21; Length 10;  
Best Local Similarity 83.3%; Pred. No. 5.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TSQYDT 10  
| | | | |  
Db 2 THQYDT 7

## RESULT 15

US-09-458-299-2  
; Sequence 2, Application US/09458299  
; GENERAL INFORMATION:  
; APPLICANT: Epimmune  
; APPLICANT: Fikes, John  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Keogh, Elissa  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO  
; FILE REFERENCE: HER2/neu USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 018623-014800US

; CURRENT APPLICATION NUMBER: US/09/458,299  
; CURRENT FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: US 09/189,702  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: US 08/205,713  
; PRIOR FILING DATE: 1994-03-04  
; PRIOR APPLICATION NUMBER: US 08/159,184  
; PRIOR FILING DATE: 1993-11-29  
; PRIOR APPLICATION NUMBER: US 08/073,205  
; PRIOR FILING DATE: 1993-06-04  
; PRIOR APPLICATION NUMBER: US 08/027,146  
; PRIOR FILING DATE: 1993-03-05  
; NUMBER OF SEQ ID NOS: 4570  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially Synthesized Peptide  
US-09-458-299-2

Query Match 43.4%; Score 26.5; DB 18; Length 8;  
Best Local Similarity 54.5%; Pred. No. 2.9e+06;  
Matches 6; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

Qy 1 LTYNTSQYDYY 11  
: | | | | |  
Db 1 VTYNT---DTF 8

Search completed: February 21, 2002, 17:22:18  
Job time: 178 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:19:35 ; Search time 13.59 Seconds  
(without alignments)  
38.262 Million cell updates/sec

Title: US-08-753-851-16

Perfect score: 61

Sequence: 1 LTYNTSQYDRY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 235865 seqs, 47271153 residues

Total number of hits satisfying chosen parameters: 45562

Minimum DB seq length: 0

Maximum DB seq length: 11

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	26	42.6	9	5 US-09-724-965-216	Sequence 216, App
2	25	41.0	10	5 US-09-572-270B-473	Sequence 473, App
3	24	39.3	11	5 US-09-824-053-7	Sequence 7, Appli
4	24	39.3	11	5 US-09-824-053-7	Sequence 14, Appl
5	23	37.7	9	5 US-09-186-949A-17	Sequence 17, Appl
6	23	37.7	9	5 US-09-620-405B-488	Sequence 488, App
7	23	37.7	9	6 US-10-007-805-488	Sequence 488, App
8	23	37.7	11	6 US-10-020-269-25	Sequence 25, Appl
9	22	36.1	9	5 US-09-674-507-5	Sequence 5, Appli
10	22	36.1	10	5 US-09-572-404B-2605	Sequence 2605, Ap
11	22	36.1	11	5 US-09-667-365-124	Sequence 124, App
12	22	36.1	11	5 US-09-667-365-348	Sequence 348, App
13	21.5	35.2	10	5 US-09-958-231-5	Sequence 5, Appli
14	21.5	35.2	11	5 US-09-667-365-156	Sequence 156, App
15	21.5	35.2	11	5 US-09-667-365-604	Sequence 604, App
16	21.5	35.2	11	5 US-09-667-365-1052	Sequence 1052, Ap
17	21.5	35.2	11	5 US-09-667-365-1500	Sequence 1500, Ap
18	21	34.4	8	5 US-09-828-708-50	Sequence 50, Appl
19	21	34.4	8	5 US-09-828-708-53	Sequence 53, Appl
20	21	34.4	8	5 US-09-483-831B-40	Sequence 40, Appl
21	21	34.4	9	1 PCT-US01-28124A-76	Sequence 76, Appl
22	21	34.4	9	5 US-09-749-873-120	Sequence 120, App
23	21	34.4	9	5 US-09-894-018-188	Sequence 188, App
24	21	34.4	9	5 US-09-948-939-36	Sequence 36, Appl
25	21	34.4	9	5 US-09-311-784A-375	Sequence 375, App
26	21	34.4	9	5 US-09-508-413A-12	Sequence 12, Appl

27 21 34.4 9 6 US-10-042-202-21 Sequence 21, Appli  
28 21 34.4 9 7 US-60-324-100-6 Sequence 6, Appli  
29 21 34.4 10 5 US-09-572-270B-471 Sequence 471, App  
30 21 34.4 11 5 US-09-692-170B-3 Sequence 3, Appli  
31 21 34.4 11 5 US-09-692-170B-4 Sequence 4, Appli  
32 21 34.4 11 5 US-09-692-170B-14 Sequence 14, Appli  
33 20 32.8 10 5 US-09-572-270B-730 Sequence 730, App  
34 20 32.8 10 5 US-09-572-270B-957 Sequence 957, App  
35 20 32.8 10 5 US-09-572-404B-951 Sequence 951, App  
36 20 32.8 10 5 US-09-572-404B-2585 Sequence 2585, Ap  
37 20 32.8 10 5 US-09-743-482A-8 Sequence 8, Appli  
38 20 32.8 10 5 US-09-743-329-1 Sequence 1, Appli  
39 20 32.8 10 5 US-09-458-302A-383 Sequence 383, App  
40 20 32.8 10 5 US-09-458-302A-1390 Sequence 1390, Ap  
41 20 32.8 11 5 US-09-667-365-138 Sequence 138, App  
42 20 32.8 11 5 US-09-667-365-139 Sequence 139, App  
43 20 32.8 11 5 US-09-667-365-140 Sequence 140, App  
44 20 32.8 11 5 US-09-667-365-141 Sequence 141, App  
45 20 32.8 11 5 US-09-667-365-572 Sequence 572, App

#### ALIGNMENTS

RESULT 1  
US-09-724-965-216  
; Sequence 216, Application US/09724965  
; GENERAL INFORMATION:  
; APPLICANT: Lonberg, Nils  
; TITLE OF INVENTION: Transgenic Non-Human Animals for  
; Producing Heterologous Antibodies  
; NUMBER OF SEQUENCES: 417  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/724,965  
; FILING DATE: 28-Nov-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/758,417  
; FILING DATE: 02-DEC-1996  
; APPLICATION NUMBER: US 08/728,463  
; FILING DATE: 10-OCT-1996  
; APPLICATION NUMBER: US 08/544,404  
; FILING DATE: 10-OCT-1995  
; APPLICATION NUMBER: US 08/352,322  
; FILING DATE: 07-DEC-1994  
; APPLICATION NUMBER: US 08/209,741  
; FILING DATE: 09-MAR-1994  
; APPLICATION NUMBER: US 08/165,699  
; FILING DATE: 10-DEC-1993  
; APPLICATION NUMBER: US 08/161,739  
; FILING DATE: 03-DEC-1993  
; APPLICATION NUMBER: US 08/155,301  
; FILING DATE: 18-NOV-1993  
; APPLICATION NUMBER: US 08/096,762  
; FILING DATE: 22-JUL-1993  
; APPLICATION NUMBER: US 08/053,131  
; FILING DATE: 26-APR-1993  
; APPLICATION NUMBER: US 07/990,860  
; FILING DATE: 16-DEC-1992  
; APPLICATION NUMBER: US 07/904,068



```
; FILING DATE: 23-JUN-1992
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; APPLICATION NUMBER: US 07/575,962
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: US 07/574,748
; FILING DATE: 29-AUG-1990
; APPLICATION NUMBER: WO PCT/US92/10983
; FILING DATE: 29-AUG-1991
; APPLICATION NUMBER: WO PCT/US91/06185
; FILING DATE: 29-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Serafini, Andrew T.
; REGISTRATION NUMBER: 41,303
; REFERENCE/DOCKET NUMBER: 014643-009030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 216:
US-09-724-965-216

Query Match 42.6%; Score 26; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 2.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 QYDTY 11
Db 2 QYDSY 6

RESULT 2
US-09-572-270B-473
; Sequence 473, Application US/09572270B
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from plant genomes
; FILE REFERENCE: Plant patent
; CURRENT APPLICATION NUMBER: US/09/572,270B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 1144
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 473
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; FEATURE:
; OTHER INFORMATION: Sequence located in Unknown at 79-88 and may interact with Sequen
US-09-572-270B-473

Query Match 41.0%; Score 25; DB 5; Length 10;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTYNTSQYDT 10
Db 1 LTYSTDDST 10

RESULT 3
US-09-824-053-7
; Sequence 7, Application US/09824053
; GENERAL INFORMATION:
; APPLICANT: Peter Stougaard
; TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hunton & Williams
; STREET: 1900 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006-1109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/824,053
; FILING DATE: 03-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/669,304
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Stanislaus Aksman
; REGISTRATION NUMBER: 28,562
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 955-1926
; TELEFAX: (202) 778-2201
; TELEX: None
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-824-053-7

Query Match 39.3%; Score 24; DB 5; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 QYDTY 11
Db 2 QYQTY 6

RESULT 4
US-09-824-053-14
; Sequence 14, Application US/09824053
; GENERAL INFORMATION:
; APPLICANT: Peter Stougaard
; TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hunton & Williams
; STREET: 1900 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006-1109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/824,053  
APPLICATION NUMBER: US/09/824,053  
FILING DATE: 03-Apr-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/669,304  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Stanislaus Aksman  
REGISTRATION NUMBER: 28,562  
REFERENCE/DOCKET NUMBER: <Unknown>  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 955-1926  
TELEFAX: (202) 778-2201  
TELEX: None  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-824-053-14

Query Match 39.3%; Score 24; DB 5; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 QYDTY 11  
|| ||  
Db 2 QYQTY 6

RESULT 5  
US-09-186-949A-17  
Sequence 17, Application US/09186949A  
GENERAL INFORMATION:  
APPLICANT: Murgita, Robert A.  
TITLE OF INVENTION: Recombinant Alpha-Fetoprotein For  
; Sequence 17, Application US/09186949A  
; FILE REFERENCE: 06727/004002  
; TITLE OF INVENTION: Treating and Diagnosing Cancers  
; CURRENT APPLICATION NUMBER: US/09/186,949A  
; CURRENT FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: US 08/758,757  
; PRIOR FILING DATE: 1996-12-03  
; PRIOR APPLICATION NUMBER: US 08/377,311  
; PRIOR FILING DATE: 1995-01-24  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-186-949A-17

Query Match 37.7%; Score 23; DB 5; Length 9;  
Best Local Similarity 55.6%; Pred. No. 2.1e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TYNTSQYDT 10  
: | | | |  
Db 1 SYICSQQDT 9

RESULT 6  
US-09-620-405B-488  
Sequence 488, Application US/09620405B  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu

APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.470C8  
; CURRENT APPLICATION NUMBER: US/09/620,405B  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 495  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 488  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-620-405B-488

Query Match 37.7%; Score 23; DB 5; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LTYNTSOY 8  
|| |||  
Db 2 LTRASQY 9

RESULT 7  
US-10-007-805-488  
Sequence 488, Application US/10007805  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Durham, Margarita  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.470C10  
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; CURRENT APPLICATION NUMBER: US/10/007,805  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 593  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 488  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-007-805-488

Query Match 37.7%; Score 23; DB 6; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LTYNTSOY 8  
|| |||  
Db 2 LTRASQY 9

RESULT 8  
US-10-020-269-25  
Sequence 25, Application US/10020269  
GENERAL INFORMATION:  
APPLICANT: Mitchell, William M.  
APPLICANT: Stratton, Charles W.  
TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE

; TITLE OF INVENTION: SEQUENCES  
; FILE REFERENCE: VDB98-01  
; CURRENT APPLICATION NUMBER: US/10/020,269  
; CURRENT FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 25  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-10-020-269-25

Query Match 37.7%; Score 23; DB 6; Length 11;  
Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 NTSQYD 9  
: || ||  
DB 3 STSQYD 8

RESULT 9  
US-09-674-507-5  
; Sequence 5, Application US/09674507  
; GENERAL INFORMATION:  
; APPLICANT: KOIBUCHI, Kyoko  
; APPLICANT: NAGASAKI, Hiroaki  
; APPLICANT: YUASA, Ari  
; APPLICANT: KATAOKA, Jiro  
; APPLICANT: KITAMOTO, Katsuhiko  
; TITLE OF INVENTION: A Novel Glutaminase, its Gene and a Method of Producing It  
; FILE REFERENCE: 199438US-B222-10-0-PCT  
; CURRENT APPLICATION NUMBER: US/09/674,507  
; CURRENT FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: JP 10/134080  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: JP 10/258974  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: JP 10/292443  
; PRIOR FILING DATE: 1998-10-14  
; PRIOR APPLICATION NUMBER: JP 11/89157  
; PRIOR FILING DATE: 1999-03-30  
; PRIOR APPLICATION NUMBER: PCT/JP99/02455  
; PRIOR FILING DATE: 1999-05-12  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Aspergillus oryzae  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: Xaa = Thr or Gly  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (8)..(8)  
; OTHER INFORMATION: Xaa is uncertain  
US-09-674-507-5

Query Match 36.1%; Score 22; DB 5; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2.1e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 QYDXY 11  
: || |  
DB 5 EYDXY 9

RESULT 10  
US-09-572-404B-2605  
; Sequence 2605, Application US/09572404B  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 2605  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; OTHER INFORMATION: sequence located in TFCOUP1 OR NR2F1 OR ERBAL3 OR EAR3 at 194-  
; OTHER INFORMATION: may interact with Sequence 2606 in this patent.  
US-09-572-404B-2605

Query Match 36.1%; Score 22; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YNTSQY 8  
: || ||  
DB 5 YPTSRY 10

RESULT 11  
US-09-667-365-124  
; Sequence 124, Application US/09667365  
; GENERAL INFORMATION:  
; APPLICANT: SUGANUMA, MASASHI  
; APPLICANT: KAWABE, TAKUMI  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING G2 CELL CYCLE  
; FILE REFERENCE: 087533/0276411  
; CURRENT APPLICATION NUMBER: US/09/667,365  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 1948  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 124  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-667-365-124

Query Match 36.1%; Score 22; DB 5; Length 11;  
Best Local Similarity 44.4%; Pred. No. 3.3e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTYNTSQYD 9  
: : || ||  
DB 1 LYHSPSQYE 9

RESULT 12  
US-09-667-365-348  
; Sequence 348, Application US/09667365  
; GENERAL INFORMATION:  
; APPLICANT: SUGANUMA, MASASHI  
; APPLICANT: KAWABE, TAKUMI  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING G2 CELL CYCLE  
; FILE REFERENCE: 087533/0276411  
; CURRENT APPLICATION NUMBER: US/09/667,365

4  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 1948  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 348  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-667-365-348

Query Match 36.1%; Score 22; DB 5; Length 11;  
Best Local Similarity 44.4%; Pred. No. 3.3e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LTYNTSQYD 9  
| : : | | |  
Db 1 LFHSPSQYE 9

## RESULT 13

US-09-958-231-5  
; Sequence 5, Application US/09958231  
; GENERAL INFORMATION:  
; APPLICANT: Sode, Koji  
; TITLE OF INVENTION: Glucose Dehydrogenase  
; FILE REFERENCE: 0230-0170P  
; CURRENT APPLICATION NUMBER: US/09/958,231  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: JP 11-101143  
; PRIOR FILING DATE: 1999-04-08  
; PRIOR APPLICATION NUMBER: JP 2000-9152  
; PRIOR FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 16

SEQ ID NO 5  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Acinetobacter calcoaceticus  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 9  
; OTHER INFORMATION: xaa is any amino acid residue  
; NAME/KEY: UNSURE  
; LOCATION: 10  
; OTHER INFORMATION: xaa is any amino acid residue  
US-09-958-231-5

Query Match 35.2%; Score 21.5; DB 5; Length 10;  
Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 2 TYNTSQYD 9  
| : : | | |  
Db 2 TYSTT-YD 8

## RESULT 14

US-09-667-365-156  
; Sequence 156, Application US/09667365  
; GENERAL INFORMATION:  
; APPLICANT: SUGANUMA, MASASHI  
; APPLICANT: KAWABE, TAKUMI  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING G2 CELL CYCLE  
; TITLE OF INVENTION: ARREST AND SENSITIZING CELLS TO DNA DAMAGING AGENTS  
; FILE REFERENCE: 087533/0276411  
; CURRENT APPLICATION NUMBER: US/09/667,365  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 1948  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 156

; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-667-365-156

Query Match 35.2%; Score 21.5; DB 5; Length 11;  
Best Local Similarity 62.5%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 3 YNT-SQYD 9  
| : : | | |  
Db 2 YNPSQYE 9

## RESULT 15

US-09-667-365-604  
; Sequence 604, Application US/09667365  
; GENERAL INFORMATION:  
; APPLICANT: SUGANUMA, MASASHI  
; APPLICANT: KAWABE, TAKUMI  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING G2 CELL CYCLE  
; TITLE OF INVENTION: ARREST AND SENSITIZING CELLS TO DNA DAMAGING AGENTS  
; FILE REFERENCE: 087533/0276411  
; CURRENT APPLICATION NUMBER: US/09/667,365  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 1948  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 604  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-667-365-604

Query Match 35.2%; Score 21.5; DB 5; Length 11;  
Best Local Similarity 62.5%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 3 YNT-SQYD 9  
| : : | | |  
Db 2 YNPSQYE 9

Search completed: February 21, 2002, 17:22:38  
Job time: 183 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2002, 17:18:30 ; Search time 12.46 Seconds  
(without alignments)  
19.866 Million cell updates/sec

Title: US-08-753-851-16  
Perfect score: 61  
Sequence: 1 LTYNTSQYDTY 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 68711

Minimum DB seq length: 0  
Maximum DB seq length: 11

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	61	100.0	11	2	US-08-143-311B-14
2	26	42.6	9	4	US-09-042-353-368
3	26	42.6	9	4	US-08-758-417A-216
4	26	42.6	11	1	US-08-190-788A-280
5	26	42.6	11	1	US-08-465-391A-280
6	26	42.6	11	2	US-08-464-538B-280
7	26	42.6	11	2	US-08-463-078E-370
8	24	39.3	10	2	US-08-637-759B-133
9	24	39.3	10	3	US-08-871-355A-133
10	24	39.3	11	4	US-08-669-304-7
11	24	39.3	11	4	US-08-669-304-14
12	23.5	38.5	11	3	US-09-120-365-28
13	23.5	38.5	11	4	US-09-515-039-28
14	23	37.7	6	6	5276016-3
15	23	37.7	9	2	US-08-377-309-16
16	23	37.7	9	4	US-09-186-723-16
17	23	37.7	9	5	PCT-US95-11127-3
18	23	37.7	9	5	PCT-US96-00996-13
19	23	37.7	10	1	US-08-267-092A-36
20	23	37.7	10	2	US-08-540-412-37
21	23	37.7	10	2	US-08-361-517-17
22	23	37.7	10	4	US-09-051-342-37
23	23	37.7	10	4	US-08-468-161-37
24	23	37.7	10	4	US-09-051-759-37
25	23	37.7	10	5	PCT-US93-07964-17
26	23	37.7	10	5	PCT-US95-08156-37
27	23	37.7	11	1	US-08-465-391A-334

28 23 37.7 11 1 US-08-511-662-5  
29 23 37.7 11 2 US-08-464-538B-332  
30 23 37.7 11 2 US-08-463-076E-371  
31 23 37.7 11 5 PCT-US96-12632-5  
32 22 36.1 9 2 US-08-722-258-67  
33 22 36.1 9 2 US-08-348-353-1  
34 22 36.1 9 2 US-08-465-965-1  
35 22 36.1 9 3 US-08-465-966-1  
36 22 36.1 9 3 US-08-599-226-18  
37 22 36.1 9 4 US-09-125-098-18  
38 22 36.1 9 5 PCT-US92-10068-18  
39 22 36.1 10 1 US-08-039-137-25  
40 22 36.1 10 3 US-08-722-258-68  
41 22 36.1 11 1 US-08-465-391A-335  
42 22 36.1 11 2 US-08-464-538B-333  
43 22 36.1 11 2 US-08-463-076E-372  
44 21 34.4 7 1 US-07-942-245-505  
45 21 34.4 8 1 US-08-346-455B-40

## ALIGNMENTS

RESULT 1  
US-08-143-311B-14  
; Sequence 14, Application US/08143311B  
; Patent No. 5863540  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, BARTON F.  
; APPLICANT: HALE, LAURA P.  
; APPLICANT: PATTON, KAREN L.  
; APPLICANT: TELEN, MARILYN J.  
; APPLICANT: LIAO, HUA-XIN  
; TITLE OF INVENTION: AN ADHESION MOLECULE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,311B  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,339  
; FILING DATE: 30-OCT-1992  
; CLASSIFICATION: 436  
; APPLICATION NUMBER: 07/669,730  
; FILING DATE: 15-MAR-1991  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-50  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-143-311B-14

Query Match 100.0%; Score 61; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 7.8e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LTYNTSQDYIY 11  
Db 1 LTYNTSQDYIY 11

RESULT 2  
US-09-042-353-368  
Sequence 368, Application US/09042353  
Patent No. 6255458  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for  
Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 421  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/042.353  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/810,279  
FILING DATE: 17-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/853,408  
FILING DATE: 18-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/904,068  
FILING DATE: 23-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/053,131  
FILING DATE: 26-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/096,762  
FILING DATE: 22-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/155,301  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/161,739  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/165,699  
FILING DATE: 10-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/209,741  
FILING DATE: 09-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/352,322  
FILING DATE: 07-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/544,404  
FILING DATE: 10-OCT-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/728,463  
FILING DATE: 10-OCT-1996  
APPLICATION NUMBER: US 08/544,404  
FILING DATE: 10-OCT-1995  
APPLICATION NUMBER: US 08/352,322  
FILING DATE: 07-DEC-1994  
APPLICATION NUMBER: US 08/209,741  
FILING DATE: 09-MAR-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/728,463  
FILING DATE: 10-OCT-1996  
APPLICATION NUMBER: US 08/544,404  
FILING DATE: 10-OCT-1995  
APPLICATION NUMBER: US 08/352,322  
FILING DATE: 07-DEC-1994  
APPLICATION NUMBER: US 08/209,741  
FILING DATE: 09-MAR-1994

Query Match 42.6%; Score 26; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 QYDTY 11  
Db 2 QYDSY 6

RESULT 3  
US-08-758-417A-216  
Sequence 216, Application US/08758417A  
Patent No. 6300129  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for  
Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 417  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/758,417A  
APPLICATION NUMBER: US/08/758,417A  
FILING DATE: 02-DEC-1996  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/728,463  
FILING DATE: 10-OCT-1996  
APPLICATION NUMBER: US 08/544,404  
FILING DATE: 10-OCT-1995  
APPLICATION NUMBER: US 08/352,322  
FILING DATE: 07-DEC-1994  
APPLICATION NUMBER: US 08/209,741  
FILING DATE: 09-MAR-1994

APPLICATION NUMBER: US 08/165,699  
FILING DATE: 10-DEC-1993  
APPLICATION NUMBER: US 08/161,739  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/155,301  
FILING DATE: 18-NOV-1993  
APPLICATION NUMBER: US 08/096,762  
FILING DATE: 22-JUL-1993  
APPLICATION NUMBER: US 08/053,131  
FILING DATE: 26-APR-1993  
APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Serafini, Andrew T.  
REGISTRATION NUMBER: 41,903  
REFERENCE/DOCKET NUMBER: 014643-0090300S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 216:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 216:  
US-08-758-417A-216

Query Match 42.6%; Score 26; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.6e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 QYDVTY 11  
| | | | |  
Db 2 QYDSY 6

RESULT 4  
US-08-190-788A-280  
Sequence 280, Application US/08190788A  
Patent No. 5608035  
GENERAL INFORMATION:  
APPLICANT: Yanofsky, Stephen D.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Baldwin, David N.  
APPLICANT: Jacobs, Jeff W.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 312  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies N.V.  
STREET: 4001 Miranda Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,788A  
FILING DATE: 02-FEB-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,567  
FILING DATE: 05-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691

REFERENCE/DOCKET NUMBER: 1019.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 280:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-190-788A-280

Query Match 42.6%; Score 26; DB 1; Length 11;  
Best Local Similarity 50.0%; Pred. No. 71;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 NTSQYDVTY 11  
| | | | |  
Db 2 NSSWYDSF 9

RESULT 5  
US-08-465-391A-280  
Sequence 280, Application US/08465391A  
Patent No. 5786331  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Yanofsky, Stephen D.  
APPLICANT: Baldwin, David  
APPLICANT: Jacobs, Jeff W.  
APPLICANT: Bovy, Philippe R.  
APPLICANT: Leahy, Ellen M.  
APPLICANT: Pottorf, Richard S.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 405  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,391A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,474  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,788  
FILING DATE: 02-FEB-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5786331v1el, Vern  
REGISTRATION NUMBER: 32,483  
REFERENCE/DOCKET NUMBER: 16528A-001840/1019.2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 280:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid



STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-391A-280

Query Match 42.6%; Score 26; DB 1; Length 11;  
Best Local Similarity 50.0%; Pred. No. 71;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 NTSQYDTY 11  
I:| | |:  
Db 2 NSSWDSF 9

## RESULT 6

US-08-464-538B-280  
Sequence 280, Application US/08464538B  
Patent No. 5861476  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Yanofsky, Stephen D.  
APPLICANT: Baldwin, David  
APPLICANT: Jacobs, Jeff W.  
APPLICANT: Bovy, Philippe R.  
APPLICANT: Leahy, Ellen M.  
APPLICANT: Pottorf, Richard S.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 402  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,538B  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,474  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,788  
FILING DATE: 02-FEB-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 16528A-001810  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 280:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-538B-280

Query Match 42.6%; Score 26; DB 2; Length 11;  
Best Local Similarity 50.0%; Pred. No. 71;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 4 NTSQYDTY 11  
I:| | |:  
Db 2 NSSWDSF 9

## RESULT 7

US-08-463-076E-370  
Sequence 370, Application US/08463076E  
Patent No. 5880096  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Yanofsky, Stephen D.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 392  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,076E  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Snyder, Joseph R.  
REGISTRATION NUMBER: 39,381  
REFERENCE/DOCKET NUMBER: 16528A-001850US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 370:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-076E-370

Query Match 42.6%; Score 26; DB 2; Length 11;  
Best Local Similarity 50.0%; Pred. No. 71;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 NTSQYDTY 11  
I:| | |:  
Db 2 NSSWDSF 9

## RESULT 8

US-08-637-759B-133  
Sequence 133, Application US/08637759B  
Patent No. 5876931  
GENERAL INFORMATION:  
APPLICANT: David William Holden  
TITLE OF INVENTION: Identification of Genes  
NUMBER OF SEQUENCES: 501  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: Georgia

COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,759B  
FILING DATE: 03-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/02875  
FILING DATE: 11-DEC-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: RPMS 101  
TELEPHONE: (404) 873-8794  
TELEFAX: (404) 873-8795  
INFORMATION FOR SEQ ID NO: 133:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-637-759B-133

Query Match 39.3%; Score 24; DB 2; Length 10;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LTYNTSQYDT 10  
| | | | |  
Db 1 LKYRTKRYSS 10

RESULT 9  
US-08-871-355A-133  
; Sequence 133, Application US/08871355A  
; Patent No. 6015669  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871,355A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.

COUNTRY: USA  
ZIP: 31,284  
REFERENCE/DOCKET NUMBER: RPMS 101 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 873-8794  
TELEFAX: (404) 873-8795  
INFORMATION FOR SEQ ID NO: 133:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-871-355A-133

Query Match 39.3%; Score 24; DB 3; Length 10;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LTYNTSQYDT 10  
| | | | |  
Db 1 LKYRTKRYSS 10

RESULT 10  
US-08-669-304-7  
; Sequence 7, Application US/08669304  
; Patent No. 6251626  
; GENERAL INFORMATION:  
; APPLICANT: Peter Stougaard  
; APPLICANT: Ole Cai Hansen  
; TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A  
; TITLE OF INVENTION: METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hulton & Williams  
; STREET: 1900 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006-1109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/669,304  
; FILING DATE: 12 July 1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/476,910  
; FILING DATE: 7 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stanislaus Aksman  
; REGISTRATION NUMBER: 28,562  
; REFERENCE/DOCKET NUMBER: 320,000003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 955-1926  
; TELEFAX: (202) 778-2201  
; TELEX: No. 6251626e  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-669-304-7

Query Match 39.3%; Score 24; DB 4; Length 11;

Best Local Similarity 80.0%; Pred. No. 1.6e+02; Mismatches 1; Indels 0; Gaps 0;

QY 7 QYDTY 11  
11111  
Db 2 QYQTY 6

RESULT 11  
US-08-669-304-14  
; Sequence 14, Application US/08669304  
; Patent No. 6251626  
; GENERAL INFORMATION:  
; APPLICANT: Peter Stougaard  
; APPLICANT: Ole Cai Hansen  
; TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A  
; TITLE OF INVENTION: METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hunton & Williams  
; STREET: 1900 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006-1109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/669,304  
; FILING DATE: 12 July 1996  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/476,910  
; FILING DATE: 7 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stanislaus Aksman  
; REGISTRATION NUMBER: 28,562  
; REFERENCE/DOCKET NUMBER: 320.000003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 955-1926  
; TELEFAX: (202) 778-2201  
; TELEX: No. 6251626g  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-669-304-14

Query Match 39.3%; Score 24; DB 4; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 QYDTY 11  
11111  
Db 2 QYQTY 6

RESULT 12  
US-09-120-365-28  
; Sequence 28, Application US/09120365  
; Patent No. 6103514  
; GENERAL INFORMATION:  
; APPLICANT: Natori, Shunji  
; TITLE OF INVENTION: NEW PROTEASE  
; FILE REFERENCE: 32290-144749  
; CURRENT APPLICATION NUMBER: US/09/120,365

; CURRENT FILING DATE: 1998-07-22  
; EARLIER APPLICATION NUMBER: JP 9-333 474  
; EARLIER FILING DATE: 1997-11-18  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 28  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Sarcophaga peregrina  
US-09-120-365-28

Query Match 38.5%; Score 23.5; DB 3; Length 11;  
Best Local Similarity 60.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 3 YNT---SOYD 9  
1111111  
Db 2 YNTLLGSHYD 11

RESULT 13  
US-09-515-039-28  
; Sequence 28, Application US/09515039  
; Patent No. 6214599  
; GENERAL INFORMATION:  
; APPLICANT: Natori, Shunji  
; TITLE OF INVENTION: NEW PROTEASE  
; FILE REFERENCE: 32290-144749  
; CURRENT APPLICATION NUMBER: US/09/515,039  
; CURRENT FILING DATE: 2000-03-06  
; EARLIER APPLICATION NUMBER: JP 9-333 474  
; EARLIER FILING DATE: 1997-11-18  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 28  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Sarcophaga peregrina  
US-09-515-039-28

Query Match 38.5%; Score 23.5; DB 4; Length 11;  
Best Local Similarity 60.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 3 YNT---SOYD 9  
1111111  
Db 2 YNTLLGSHYD 11

RESULT 14  
5276016-3  
; Patent No. 5276016  
; APPLICANT: PERT, CANDACE B.; RUFF, WILLIAM R.  
; FARRAR, WILLIAM L.  
; TITLE OF INVENTION: SMALL PEPTIDES WHICH INHIBIT  
; BINDING TO T-4 RECEPTORS AND ACT AS IMMUNOGENS  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/568,616  
; FILING DATE: 16-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 314,507  
; FILING DATE: 15-FEB-1989  
; APPLICATION NUMBER: 48,148  
; FILING DATE: 11-MAY-1987  
; APPLICATION NUMBER: 878,586  
; FILING DATE: 26-JUN-1986  
; APPLICATION NUMBER: 869,919  
; FILING DATE: 03-JUN-1986  
; SEQ ID NO:3:  
; LENGTH: 6

52760183

Query Match 37.7%; Score 23; DB 6; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.6e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TSQYDT 10  
I |||  
Db 1 TDNYDT 6

RESULT 15  
US-08-377-309-16  
; Sequence 16, Application US/08377309A  
; Patent No. 5965528  
; GENERAL INFORMATION:  
; APPLICANT: Murgita, Robert A.  
; TITLE OF INVENTION: RECOMBINANT HUMAN ALPHA-FETOPROTEIN AS  
; FILE REFERENCE: 06727/005001  
; CURRENT APPLICATION NUMBER: US/08/377,309A  
; CURRENT FILING DATE: 1995-01-24  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-377-309-16

Query Match 37.7%; Score 23; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.6e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 TYNTSQYDT 10  
: |||||  
Db 1 SYICSQQDT 9

Search completed: February 21, 2002, 17:20:21  
Job time: 111 sec